Note di Algebra Lineare

Nicoletta Cantarini[1]

[1]Liberamente (es)tratto da: Un Corso di Matematica, N. Cantarini, B. Chiarellotto, L. Fiorot, Ed. Progetto, Padova 2006

Indice

Introduzione

Queste note non hanno la pretesa di sostituirsi ad uno dei numerosi testi di Algebra Lineare disponibili in letteratura ma semplicemente di offrire agli studenti del corso di Laurea in Informatica per il Management dell'Università di Bologna un supporto nella preparazione dell'esame di Algebra Lineare.

Descriviamo un paio di problemi che gli studenti saranno in grado di risolvere alla fine del corso.

Problema A. (Problema enigmistico di basso livello)

i) Calcolare le età di due sorelle, età che indicheremo con E_1 ed E_2, sapendo che l'età della prima sommata a 2 volte l'età della seconda è pari a 22 e che 3 volte l'età della prima meno 2 volte l'età della seconda è pari a 2. Risolvere tale problema significa trovare E_1 ed E_2 tali che le due equazioni: $E_1+2E_2 = 22$ e $3E_1-2E_2 = 2$ siano soddisfatte contemporaneamente. Dalla prima equazione si ottiene $E_1 = 22 - 2E_2$ e, sostituendo questa espressione nella seconda equazione, si ottiene $E_2 = 8$ da cui $E_1 = 6$.

Potremmo rendere le cose più complicate facendo entrare in gioco anche l'età di una zia che indichiamo con Z. Allora il quesito potrebbe essere il seguente: calcolare le tre età E_1, E_2, Z, sapendo che l'età della prima sorella meno l'età della seconda meno quella della zia è pari a 2, e che 2 volte l'età della zia meno l'età della prima sorella più l'età della seconda è pari a 4. Allora $Z = 6$, $E_2 = 2$ ed $E_1 = 10$ è una soluzione, ma anche $Z = 6$, $E_2 = 4$ ed $E_1 = 12$ lo è. Quindi tali problemi possono avere diverse soluzioni, ma quante esattamente? Quando possiamo affermare con certezza che il problema ha una sola soluzione, come nel primo caso?

ii) Un secondo esempio di applicazione dei sistemi lineari viene dalla chimica. Supponiamo di voler bilanciare un'equazione chimica. Ad esempio, consideriamo la reazione tra sale comune NaCl e acido sulfureo $\mathrm{H_2SO_4}$:

$$\mathrm{NaCl} + \mathrm{H_2SO_4} \to \mathrm{Na_2SO_4} + \mathrm{HCl}.$$

È immediato vedere che per bilanciare tale equazione si trova

$$2\mathrm{NaCl} + \mathrm{H_2SO_4} \to \mathrm{Na_2SO_4} + 2\mathrm{HCl}$$

Bilanciare un'equazione chimica equivale a richiedere che il numero di atomi di ogni elemento prima di una reazione sia pari al numero di atomi presente dopo la reazione. Quindi per bilanciare l'equazione

$$x\mathrm{C_7H_8} + y\mathrm{HNO_3} \to z\mathrm{C_7H_5O_6N_3} + w\mathrm{H_2O}$$

dovremo risolvere il sistema lineare

$$\begin{cases} 7x = 7z \\ 8x + 1y = 5z + 2w \\ 1y = 3z \\ 3y = 6z + 1w. \end{cases}$$

Problema B. (Evoluzione del sistema)

Supponiamo che in un'isola vi siano volpi in numero pari a V e galline in numero pari a G. Supponiamo sia dato un modello per cui in un anno le galline si riproducono determinando un aumento della popolazione del 60 per cento mentre le volpi si mangiano le galline per un fattore 1 rispetto al loro numero. Come si sarà evoluto il sistema dopo un anno? Il numero di galline, che indichiamo con G_1, sarà pari a $G_1 = 1,6G_0 - V_0$ ovvero al numero iniziale di galline G_0 a cui si è aggiunto il numero di pulcini $0,6G_0$ meno il numero di galline mangiate dalle volpi, pari al numero iniziale di volpi, cioè V_0. D'altro canto supponiamo che il tasso di natalità delle volpi sia del 10 per cento e che le galline abbiano una malattia che si trasmette alle volpi che se le mangiano in modo tale che la mortalità delle volpi a causa di questa malattia sia proporzionale a metà del numero di galline. Questo significa che dopo un anno il numero di volpi V_1 sarà pari a $V_1 = 1,1V_0 - 0,5G_0$ (dove $0,5G_0$ è la quantità di volpi uccise dalla malattia). Cosa potrebbe succedere a questo punto? Se le volpi fossero troppe alla fine si mangerebbero tutte le galline e non resterebbe più nulla da mangiare per le volpi, così nell'isola non vi sarebbe più nessuno. Quante galline ci vogliono e quante volpi occorrono per avere un sistema che non si esaurisca? Oppure, in tale situazione, per ogni scelta iniziale di galline e volpi alla fine l'isola rimarrà deserta? Ovviamente bisognerebbe conoscere a priori l'evoluzione del nostro sistema, cioè sapere a priori quello che avverrà.

Lezione 1

Spazi vettoriali

Questa lezione è dedicata allo studio degli spazi vettoriali. Attraverso il concetto di spazio vettoriale vogliamo innanzitutto costruire un modello di uno spazio di dimensione qualsiasi. Non dobbiamo dimenticare che, aldilà di qualsiasi astrazione, ognuno di noi possiede una idea intuitiva di dimensione legata alla vita quotidiana: viviamo e ci muoviamo in uno spazio (fisico) tridimensionale, disegnamo su fogli essenzialmente bidimensionali, e così via.

1.1 Definizione di spazio vettoriale reale

Cos'è uno spazio vettoriale (reale)? Uno spazio vettoriale è un insieme non vuoto V dotato di una somma, su cui 'agisce dall'esterno' l'insieme $\mathbb{R}$ dei numeri reali. Ma cosa vuol dire agire dall'esterno? Vuol dire che, preso un elemento v di questo insieme V e preso un qualsiasi elemento c di $\mathbb{R}$, viene associato a questa coppia un elemento di V che denoteremo con cv. Passiamo alla definizione vera e propria.

Definizione 1.1.1 *Diremo che un insieme non vuoto V è uno spazio vettoriale su $\mathbb{R}$ o, equivalentemente, un $\mathbb{R}$-spazio vettoriale, se:*

- *su V è definita una operazione detta somma, $+_V$ (operazione interna), che è: commutativa, associativa, ammette elemento neutro, $\mathbf{0}_V$, detto vettore nullo, e tale che ogni elemento v di V ammetta opposto denotato con $-v$;*

- *è definita un'azione esterna di $\mathbb{R}$ su V, cioè un'applicazione $\mathbb{R} \times V \to V$, che ad ogni coppia (r, v) con $r \in \mathbb{R}$ e $v \in V$ associa un unico elemento (che denoteremo con) $rv \in V$. Questa operazione è detta operazione esterna.*

Le precedenti operazioni godono delle seguenti proprietà di compatibilità:
(i) $1v = v$, *per ogni* $v \in V$.
(ii) $\forall v \in V$ *e* $\forall \alpha, \beta \in \mathbb{R}$, $(\alpha + \beta)v = \alpha v +_V \beta v$.
(iii) $\forall v, w \in V$ *e* $\forall \alpha \in \mathbb{R}$, $\alpha(v +_V w) = \alpha v +_V \alpha w$.
(iv) $\forall v \in V$ *e* $\forall \alpha, \beta \in \mathbb{R}$, $(\alpha\beta)v = \alpha(\beta v)$.

Gli elementi di V sono detti vettori, gli elementi di $\mathbb{R}$ scalari; l'operazione esterna si dice anche prodotto per scalari. Talvolta, qualora non vi sia pericolo di confusione, indicheremo la somma $+_V$ in V semplicemente con $+$. Anche se lavoreremo sempre con i numeri reali è importante sapere che nella definizione di spazio vettoriale che abbiamo appena dato, l'insieme dei numeri reali può essere sostituito con un altro insieme avente analoghe proprietà (campo), ad esempio $\mathbb{Q}$ (campo dei numeri razionali) o $\mathbb{C}$ (campo dei numeri complessi).

1.2 Proprietà degli spazi vettoriali

Elenchiamo in questo paragrafo alcune utili proprietà degli spazi vettoriali.

1. Calcoliamo il prodotto del numero reale 0 per il vettore $v \in V$: $0v$. Chi è questo elemento? Per la proprietà *(ii)* della definizione 1.1.1 si ha $0v = (0 + 0)v = 0v +_V 0v$. Sommando a destra e a sinistra l'opposto di $0v$, che denoteremo con $-0v$, si ha: $\mathbf{0}_V = 0v +_V (-0v) = 0v +_V 0v +_V (-0v) = 0v$; la prima uguaglianza vale poiché $(-0v)$ è l'opposto di $0v$, la seconda perché si aggiunge a due elementi uguali lo stesso elemento e infine l'ultima perché $0v +_V (-0v) = \mathbf{0}_V$ e $0v +_V \mathbf{0}_V = 0v$. Per la proprietà transitiva dell'uguaglianza si ha $0v = \mathbf{0}_V$ vettore nullo di V.

2. Per la definizione 1.1.1 (i) $1v = v$ per ogni $v \in V$. Ora considerando $(1+(-1))v$, per la proprietà distributiva si ottiene $\mathbf{0}_V = 0v = 1v +_V (-1)v = v +_V (-1)v$ quindi $(-1)v = -v$ è l'opposto di v. Tale opposto verrà indicato semplicemente con $-v$. Notiamo che se nel precedente ragionamento avessimo sostituito 1 con un qualsiasi numero reale α avremmo ottenuto che l'opposto di αv è $(-\alpha)v = -\alpha v$.

3. Per ogni $\alpha \in \mathbb{R}$ vale $\alpha\mathbf{0}_V = \mathbf{0}_V$. Infatti dalle proprietà precedenti degli spazi vettoriali si ha che:

$$\alpha\mathbf{0}_V = \alpha(\mathbf{0}_V +_V \mathbf{0}_V) = \alpha\mathbf{0}_V +_V \alpha\mathbf{0}_V.$$

Sommando ad ambedue i membri dell'uguaglianza l'opposto di $\alpha\mathbf{0}_V$, che indichiamo con $-\alpha\mathbf{0}_V$, otteniamo

$$\alpha\mathbf{0}_V +_V (-\alpha\mathbf{0}_V) = (\alpha\mathbf{0}_V +_V \alpha\mathbf{0}_V) +_V (-\alpha\mathbf{0}_V)$$

che, per la proprietà associativa e per la definizione di elemento neutro, equivale a

$$\mathbf{0}_V = \alpha\mathbf{0}_V + (\alpha\mathbf{0}_V +_V -\alpha\mathbf{0}_V) = \alpha\mathbf{0}_V +_V \mathbf{0}_V = \alpha\mathbf{0}_V.$$

4. Se $\alpha \neq \beta$ con $\alpha, \beta \in \mathbb{R}$, allora $\alpha v \neq \beta v$ per ogni $v \in V$, $v \neq \mathbf{0}_V$. Infatti se $\alpha v = \beta v$, allora sommando a destra e a sinistra per l'opposto di βv, si avrebbe $(\alpha - \beta)v = \mathbf{0}_V$. Da cui moltiplicando ambedue i membri per $1/(\alpha - \beta)$ (essendo $\alpha - \beta \neq 0$) si otterrebbe $v = 1/(\alpha - \beta)\mathbf{0}_V = \mathbf{0}_V$ che contraddirebbe l'ipotesi $v \neq \mathbf{0}_V$. Dunque è stato assurdo pensare che αv e βv fossero eguali.

5. Può esistere uno spazio vettoriale su $\mathbb{R}$ con un numero finito di vettori? Certamente un esempio è lo spazio vettoriale banale, $V = \{\mathbf{0}_V\}$, cioè lo spazio vettoriale costituito dal solo elemento nullo in cui le operazioni interna ed esterna sono banali: $\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$, $\alpha\mathbf{0}_V = \mathbf{0}_V$ per ogni α in $\mathbb{R}$. È facile verificare la validità degli assiomi della definizione 1.1.1. La domanda che ci poniamo ora è la seguente: esistono altri esempi di spazi vettoriali reali con un numero finito di elementi? Supponiamo che V sia uno spazio vettoriale su $\mathbb{R}$ con un numero finito di elementi, i.e. un numero finito di vettori, diciamo $n \in \mathbb{N}$. Sia $v \in V$, $v \neq \mathbf{0}_V$. Per la proprietà precedente $v, 2v, 3v, 4v, \ldots, nv$ sono tutti diversi fra loro e poiché sono n devono essere tutti e soli i vettori di V. Ne segue che, essendo $(n+1)v$ un vettore di V, $(n+1)v$ dovrà essere uguale ad uno dei precedenti, ma ciò è assurdo perché si avrebbe $(n+1)v = sv$ con $s \in \mathbb{N}$, $s \leq n$, cioè $s \neq n+1$. È stato assurdo pensare che V contenesse solo un numero finito di elementi. Quindi l'unico spazio vettoriale su $\mathbb{R}$ con un numero finito di vettori è lo spazio vettoriale banale. Il metodo appena illustrato è detto delle "gabbie dei piccioni": se si ha un numero di gabbie inferiore al numero di piccioni e se si vogliono mettere tutti i piccioni in gabbia, allora necessariamente almeno una gabbia dovrà alloggiare più di un piccione.

1.3 Esempi

1.3.1 Spazio vettoriale banale

Ogni spazio vettoriale contiene il vettore nullo cioè l'elemento neutro rispetto alla somma. Prendiamo l'insieme formato dal solo vettore nullo $\{\mathbf{0}_V\}$, cioè lo spazio vettoriale banale introdotto in 1.2.5. In esso la somma e il prodotto per scalari sono definiti in modo ovvio:

$$\mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V; \qquad \lambda\mathbf{0}_V = \mathbf{0}_V, \quad \forall\ \lambda \in \mathbb{R}.$$

A titolo di esempio si noti che, con le definizioni date, per ogni $\alpha, \beta \in \mathbb{R}$ $\mathbf{0}_V = (\alpha+\beta)\mathbf{0}_V = \alpha\mathbf{0}_V + \beta\mathbf{0}_V$. Torneremo ancora su questo esempio.

1.3.2 Spazi vettoriali $\mathbb{R}^n$

Consideriamo l'insieme delle coppie ordinate di numeri reali. (Ricordiamo che ordinare le coppie significa che in generale l'elemento (a, b) è diverso dall'elemento (b, a), ad esempio $(2, 3) \neq (3, 2)$.) Tale insieme viene indicato con $\mathbb{R}^2 = \{(\alpha_1, \alpha_2) \mid \alpha_1, \alpha_2 \in \mathbb{R}\}$. Due elementi (α_1, α_2) e (β_1, β_2) di $\mathbb{R}^2$ sono uguali se e solo se $\alpha_1 = \beta_1$ e $\alpha_2 = \beta_2$. Vogliamo definire su $\mathbb{R}^2$ una struttura di spazio vettoriale. Occorre innanzitutto introdurre un'operazione interna che definiamo "componente per componente" nel modo seguente:

$$(\alpha_1, \alpha_2) +_{\mathbb{R}^2} (\beta_1, \beta_2) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2)$$

dove la somma in ogni componente è la somma di numeri reali. L'operazione $+_{\mathbb{R}^2}$ appena definita è commutativa e associativa poiché essa viene definita componente per componente mediante una operazione (la somma di numeri reali) che gode di tali proprietà. L'elemento neutro è $\mathbf{0}_{\mathbb{R}^2} = (0, 0)$. Definiamo ora il prodotto per scalari: se $\lambda \in \mathbb{R}$ e $(\alpha_1, \alpha_2) \in \mathbb{R}^2$ definiamo $\lambda(\alpha_1, \alpha_2) = (\lambda\alpha_1, \lambda\alpha_2)$. Anche in questo caso è chiaro che tutti gli assiomi della definizione di spazio vettoriale sono rispettati.

Analogamente indicheremo con $\mathbb{R}^3$ l'insieme delle terne ordinate di numeri reali e, più in generale, con $\mathbb{R}^n$, con $n \in \mathbb{N}$, l'insieme delle $n-uple$ ordinate di numeri reali. Generalizzando quanto descritto sopra, definiamo su $\mathbb{R}^n$ una struttura di spazio vettoriale reale definendo la somma ed il prodotto per scalari come segue:

$$(\alpha_1, \alpha_2, \ldots, \alpha_n) +_{\mathbb{R}^n} (\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$$

e
$$\lambda(\alpha_1, \alpha_2, \dots, \alpha_n) = (\lambda\alpha_1, \lambda\alpha_2, \dots, \lambda\alpha_n).$$
Notiamo che tra gli spazi vettoriali così definiti vi è anche $\mathbb{R}^1 = \mathbb{R}$: l'insieme dei numeri reali inteso come spazio vettoriale su se stesso. In tal caso denotiamo i vettori di $\mathbb{R}^1$ con (a), $a \in \mathbb{R}$.

Osservazione. Il campo $\mathbb{C}$ dei numeri complessi può essere visto come spazio vettoriale reale con le usuali operazioni di somma di numeri complessi e prodotto di numeri reali con numeri complessi.

1.3.3 Matrici

Una **matrice** $n \times m$ $(n, m \in \mathbb{N})$ ad entrate in $\mathbb{R}$ è il dato di nm numeri reali a_{ij} dove $i \in \{1, 2, \dots, n\}$ e $j \in \{1, 2, \dots, m\}$, collocati su una tabella con n righe e m colonne. Il numero i indica la riga ove è posizionato l'elemento a_{ij} e j ne indica la colonna. Ad esempio la scrittura:
$$A = \begin{pmatrix} 2 & 5 & 0 \\ -1 & \sqrt{2} & 4 \end{pmatrix}$$
individua una matrice 2×3. L'elemento a_{12} è 5, l'elemento a_{23} è 4. Indichiamo l'insieme delle matrici a n righe e m colonne con $\mathcal{M}_{n,m}(\mathbb{R})$. Fissati $n, m \in \mathbb{R}$ l'insieme $\mathcal{M}_{n,m}(\mathbb{R})$, può essere munito di una struttura di spazio vettoriale. Cominciamo con l'esempio di $\mathcal{M}_{2,3}(\mathbb{R})$, gli altri casi sono del tutto analoghi e si differenziano solo per la forma delle matrici. Definiamo l'operazione interna componente per componente:
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} +_{\mathcal{M}_{2,3}(\mathbb{R})} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix}$$
dove $c_{11} = a_{11} + b_{11}$ come somma di numeri reali, $c_{12} = a_{12} + b_{12}$ e così via $c_{ij} = a_{ij} + b_{ij}$. L'elemento neutro rispetto alla somma così definita è la matrice nulla
$$\mathbf{0}_{\mathcal{M}_{2,3}(\mathbb{R})} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$
Definiamo l'operazione esterna sempre componente per componente: dati $\lambda \in \mathbb{R}$ ed $A = (a_{ij}) \in \mathcal{M}_{2,3}(\mathbb{R})$ poniamo
$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \lambda a_{13} \\ \lambda a_{21} & \lambda a_{22} & \lambda a_{23} \end{pmatrix}.$$

Usando ancora una volta le proprietà dei numeri reali si verifica facilmente che le operazioni introdotte definiscono su $\mathcal{M}_{2,3}(\mathbb{R})$ una struttura di spazio vettoriale reale. Si noti che, sia a livello di insiemi che a livello di spazi vettoriali, è possibile identificare $\mathcal{M}_{1,3}(\mathbb{R})$ con $\mathbb{R}^3$ e più in generale $\mathcal{M}_{1,n}(\mathbb{R})$ con $\mathbb{R}^n$, per ogni $n \in \mathbb{N}$. Con una piccola rotazione... della testa... si potrebbe pure identificare $\mathcal{M}_{n,1}(\mathbb{R})$ con $\mathbb{R}^n$, per ogni $n \in \mathbb{N}$, e in effetti questo è vero anche a livello di spazi vettoriali. Preciseremo meglio in seguito che cosa intendiamo.

In generale, date $A = (a_{ij})$ e $B = (b_{ij})$ matrici $n \times m$ definiamo $A + B = C$ dove $C = (c_{ij})$ è la matrice $n \times m$ di termine generico $c_{ij} = a_{ij} + b_{ij}$; inoltre, per ogni numero reale λ, definiamo λA come la matrice $n \times m$ ottenuta moltiplicando ogni entrata di A per λ:

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \dots & \dots & \dots & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}.$$

In questo modo abbiamo dotato $\mathcal{M}_{n,m}(\mathbb{R})$ di una struttura di $\mathbb{R}$-spazio vettoriale.

Se $n = m$ diremo che la matrice è quadrata di ordine n. L'insieme delle matrici quadrate di ordine n ad entrate reali si denota con $\mathcal{M}_n(\mathbb{R})$.

1.3.4 Insiemi di polinomi

Sia $\mathbb{R}[X]$ l'insieme dei polinomi nella indeterminata X a coefficienti in $\mathbb{R}$ e indichiamo con $\mathbb{R}^{\leq n}[X]$ i polinomi di $\mathbb{R}[X]$ di grado minore o uguale a n.

Anche su $\mathbb{R}[X]$ si può definire una struttura di $\mathbb{R}$-spazio vettoriale: presi due polinomi $P(X)$ e $Q(X)$, di grado rispettivamente n e m, con $n, m \in \mathbb{N}$ e $n \leq m$: $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$, $Q(X) = b_mX^m + b_{m-1}X^{m-1} + \ldots + b_1X + b_0$, definiamo la loro somma in $\mathbb{R}[X]$ come

$$P(X) +_{\mathbb{R}[X]} Q(X) = \sum_{i=0}^{m} c_i X^i$$

con $c_i = b_i$ se $m \geq i \geq n + 1$ e $c_i = b_i + a_i$ per $i \leq n$. Ad esempio, se $Q(X) = -X^2 + 4X - \sqrt{3}$ e $P(X) = 3X^4 + 2X + 6$, $P(X) +_{\mathbb{R}[X]} Q(X) = 3X^4 - X^2 + 6X + (6 - \sqrt{3})$.

La somma di polinomi così definita è commutativa e associativa; l'elemento neutro $\mathbf{0}_{\mathbb{R}[X]}$ per la somma è il polinomio di grado 0 il cui termine noto è zero, cioè il polinomio identicamente nullo. Ogni polinomio ammette opposto per tale somma: se $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$, allora il suo opposto è dato da $-P(X) = -a_nX^n - a_{n-1}X^{n-1} - \ldots - a_1X - a_0$. Sia λ un numero reale e sia $P(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ un polinomio, definiamo $\lambda P(X) = (\lambda a_n)X^n + (\lambda a_{n-1})X^{n-1} + \ldots + (\lambda a_1)X + (\lambda a_0)$, cioè moltiplichiamo ogni monomio di $P(X)$ per λ. Si vede facilmente che il prodotto per scalari così definito gode della proprietà distributiva rispetto alla somma, e che $1P(X) = P(X)$. L'insieme $\mathbb{R}[X]$ dotato delle operazioni sopra descritte è quindi uno spazio vettoriale reale. In particolare $(-1)P(X)$ è l'opposto di $P(X)$.

Osservazione. Consideriamo ora il sottoinsieme $\mathbb{R}^{\leq n}[X]$ di $\mathbb{R}[X]$. Le operazioni che abbiamo definito su $\mathbb{R}[X]$ sono definite anche in $\mathbb{R}^{\leq n}[X]$ e osserviamo che la somma di due polinomi di grado minore o uguale a n è ancora un polinomio di grado minore o uguale a n e se $\lambda \in \mathbb{R}$ e $P(X) \in \mathbb{R}^{\leq n}[X]$, allora il polinomio $\lambda P(X)$ appartiene a $\mathbb{R}^{\leq n}[X]$. In altre parole $\mathbb{R}^{\leq n}[X]$ è un sottoinsieme di $\mathbb{R}[X]$ chiuso rispetto alle operazioni di $\mathbb{R}[X]$. Dunque $\mathbb{R}^{\leq n}[X]$ ha, con queste operazioni, una struttura di spazio vettoriale.

Questo è un esempio di sottoinsieme di vettori di uno spazio vettoriale su cui le operazioni dello spazio ambiente inducono una struttura di spazio vettoriale. Un tale sottoinsieme si dice sottospazio vettoriale, ma vedremo in dettaglio questa definizione nella Lezione 2.

Notiamo infine che le nostre conoscenze matematiche ci permettono di definire un'altra operazione all'interno dell'insieme dei polinomi $\mathbb{R}[X]$: la moltiplicazione di due polinomi (i.e. $(3X^2 + 2)(5X - 1) = 15X^3 - 3X^2 + 10X - 2$). Tale operazione non interviene tuttavia nella definizione di spazio vettoriale.

1.3.5 Spazi di funzioni

Costruiamo infine un ultimo esempio di spazio vettoriale. Consideriamo l'insieme delle applicazioni continue dall'intervallo $[0, 1] \subset \mathbb{R}$ verso $\mathbb{R}$ e indichiamo tale insieme con $\mathcal{C}_{\mathbb{R}}([0, 1])$. Muniamo tale insieme di una struttura di spazio vettoriale. Dati f, g due elementi di $\mathcal{C}_{\mathbb{R}}([0, 1])$ definiamo la funzione $h \in \mathcal{C}_{\mathbb{R}}([0, 1])$ somma di f e g

$$f +_{\mathcal{C}_{\mathbb{R}}([0,1])} g = h$$

come l'applicazione da $[0,1]$ a valori in $\mathbb{R}$ tale che il valore di h in $s \in [0,1]$ sia dato da $h(s) = f(s) +_{\mathbb{R}} g(s)$, cioè dalla somma in $\mathbb{R}$ dei valori delle due funzioni f e g in s. Essendo la somma una applicazione continua allora anche h è una applicazione continua e quindi appartiene a $\mathcal{C}_{\mathbb{R}}([0,1])$. Tale somma è commutativa e associativa (prese tre funzioni f_1, f_2, f_3 di $\mathcal{C}_{\mathbb{R}}([0,1])$, allora

$$(f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_2) +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_3 = f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} (f_2 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_3)$$

poiché il valore in $s \in [0,1]$ di ciascuna somma è: $(f_1(s) + f_2(s)) + f_3(s) = f_1(s) + (f_2(s) + f_3(s))$ per l'associatività della somma di numeri reali. L'elemento neutro rispetto alla somma è la funzione $\mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}$, cioè la funzione identicamente uguale a 0; ogni $f \in \mathcal{C}_{\mathbb{R}}([0,1])$ ammette opposto, cioè esiste $g \in \mathcal{C}_{\mathbb{R}}([0,1])$ tale che

$$f +_{\mathcal{C}_{\mathbb{R}}([0,1])} g = \mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}.$$

Infatti, essendo $\mathbf{0}_{\mathcal{C}_{\mathbb{R}}([0,1])}$ la funzione identicamente nulla, si ha che l'opposto di una funzione f è la funzione g tale che $g(s) = -f(s)$ per $s \in [0,1]$. Dati $\lambda \in \mathbb{R}$ e $f \in \mathcal{C}_{\mathbb{R}}([0,1])$ definiamo $(\lambda f)(s) = \lambda f(s)$. Tale applicazione è continua, quindi λf è ancora un elemento di $\mathcal{C}_{\mathbb{R}}([0,1])$. Valutando di volta in volta le funzioni nei punti $s \in [0,1]$, si vede che le proprietà di spazio vettoriale sono verificate. Si noti ad esempio che la proprietà distributiva del prodotto per scalari rispetto alla somma è data dalla uguaglianza delle due funzioni continue

$$\lambda(f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} f_2) = \lambda f_1 +_{\mathcal{C}_{\mathbb{R}}([0,1])} \lambda f_2,$$

il che si verifica valutando le due funzioni in $s \in [0,1]$:

$$\lambda(f_1(s) + f_2(s)) = \lambda f_1(s) + \lambda f_2(s).$$

In conclusione la definizione astratta di spazio vettoriale è un formalismo che si adatta a molte situazioni in apparenza ben diverse: numeri reali, n-uple, matrici, polinomi ... Quindi non ci interessano necessariamente quali siano le operazioni introdotte per definire una struttura di spazio vettoriale perché il comportamento di ogni spazio vettoriale risulta sempre lo stesso. Ad esempio in 1.3.5 il vettore opposto a f, che sappiamo essere $(-1)f$, è $-f$. Così come in $\mathbb{R}^3$ il vettore opposto a $v = (1,-1,5)$, e formalmente dato da $(-1)v$, si determina attraverso la legge di spazio vettoriale data su $\mathbb{R}^3$: $(-1)(1,-1,5) = (-1,1,-5)$.

Lezione 2

Combinazioni lineari e sottospazi

2.1 Sottospazi vettoriali

In ogni costruzione matematica, dopo aver definito una classe di oggetti aventi delle proprietà omogenee, risulta naturale introdurre la nozione di sottooggetti che rispettino tali proprietà. Da un punto di vista insiemistico, cioè quando gli oggetti sono insiemi, un sottooggetto di un insieme W è un sottoinsieme U, cioè un insieme i cui elementi sono elementi di W: $U \subseteq W$. Vogliamo allora che nella definizione di sottooggetto di uno spazio vettoriale si tenga conto della sua struttura, più ricca di quella di un semplice insieme: dovremo dunque fare in modo che ogni sottooggetto erediti questa struttura.

Definizione 2.1.1 *Dato un $\mathbb{R}$-spazio vettoriale V, un sottospazio vettoriale U di V è un sottoinsieme non vuoto $U \subseteq V$ su cui le operazioni di V inducono una struttura di spazio vettoriale reale.*

Che cosa vuol dire esattamente questa definizione? Sia V uno spazio vettoriale reale e sia U un suo sottoinsieme non vuoto. Dati due vettori u e v di U possiamo considerare $u +_V v$ dal momento che $U \subseteq V$ e quindi gli elementi di U sono, in particolare, vettori di V. In generale il vettore somma $u +_V v$ NON sarà un elemento di U. Dire che le operazioni di V inducono una struttura di spazio vettoriale su U significa che la somma in V ed il prodotto per scalari:

$$+_V : V \times V \longrightarrow V \qquad \cdot : \mathbb{R} \times V \longrightarrow V$$

si restringono ad operazioni su U:

$$+_V : U \times U \longrightarrow U \qquad \cdot : \mathbb{R} \times U \longrightarrow U.$$

In particolare il vettore $u +_V v$ appartiene ancora ad U. In questo caso si usa dire che U è *chiuso* rispetto alla somma di V. Analogamente, dati $\lambda \in \mathbb{R}$ e $v \in U$, il vettore λv è un elemento di V ma, in generale, se U è solo un sottoinsieme di V, λu NON è un elemento di U. Se, però, U è un sottospazio vettoriale di V allora λu appartiene ancora ad U. Diremo in questo caso che U è *chiuso* rispetto al prodotto per scalari di V.

Il problema è adesso il seguente: dati un $\mathbb{R}$-spazio vettoriale V ed un suo sottoinsieme U, come facciamo a stabilire se U è un sottospazio di V? Basterà verificare che U è chiuso rispetto alle operazioni di V, usando il criterio seguente:

Proposizione 2.1.2 *Un sottoinsieme U di un $\mathbb{R}$-spazio vettoriale V è un sottospazio vettoriale di V se e solo se per ogni elemento $u \in U$ e per ogni $\lambda \in \mathbb{R}$, $\lambda u \in U$ e per ogni $u, v \in U$ la somma $u +_V v \in U$.*

Dimostrazione.

"$\Rightarrow$" Se U è un sottospazio vettoriale le operazioni di V inducono su U una struttura di spazio vettoriale, quindi le due proprietà dell'enunciato sono verificate.

"$\Leftarrow$" Viceversa supponiamo che U soddisfi le due proprietà dell'enunciato, cioè che presi due elementi in U la loro somma sia ancora in U, e per ogni elemento u di U e per ogni $\lambda \in \mathbb{R}$, $\lambda u \in U$. Dobbiamo verificare la validità degli assiomi della definizione 1.1.1. Innanzitutto la somma in U è associativa e commutativa perché lo è la somma di V. Inoltre ogni elemento u di U ha opposto in U, in quanto, per le proprietà degli spazi vettoriali, l'opposto di u è $-u = (-1)u$ che appartiene ad U per ipotesi ($\lambda = -1$). Abbiamo perciò verificato la validità di tutti gli assiomi che riguardano l'operazione interna. Analogamente possiamo ragionare per l'operazione esterna che risulta ben definita su U per ipotesi. Gli assiomi *(i),..., (iv)* della definizione 1.1.1, cioè quelli che caratterizzano le operazioni di uno spazio vettoriale, sono automaticamente verificati in U in quanto le operazioni di U sono le stesse di V. **C. V. D.**

Dato un $\mathbb{R}$-spazio vettoriale V per indicare che un sottoinsieme U di V è un suo sottospazio scriveremo $U \leq V$.

Osservazione 2.1.3 Ogni spazio vettoriale V ha sempre almeno due sottospazi: V stesso e il sottospazio vettoriale banale $\{\mathbf{0}_V\}$. Chiaramente ogni sottospazio vettoriale contiene il vettore nullo $\mathbf{0}_V$.

Vediamo ora in alcuni esempi come applicare la proposizione 2.1.2.

Esempi 2.1.4 Supponiamo che il nostro spazio ambiente sia $V = \mathbb{R}^2$:

1. Consideriamo l'insieme costituito dal solo vettore nullo: $\{(0,0)\}$. Esso è un sottospazio vettoriale di $\mathbb{R}^2$: il sottospazio banale.

2. Sia $U = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. L'insieme U non può essere un sottospazio vettoriale poiché non contiene il vettore nullo di $\mathbb{R}^2$ che è $(0,0)$.

3. Cerchiamo di sopperire a questo fatto considerando ora $U = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$, che è il cerchio di centro l'origine e raggio 1. Questa volta $(0,0) \in U$, ma U non è chiuso rispetto a $+_{\mathbb{R}^2}$: infatti, presi, ad esempio, $(1,0), (0,1) \in U$, la loro somma $(1,0) +_{\mathbb{R}^2} (0,1) = (1,1)$ non è un elemento del cerchio, quindi U non è un sottospazio di $\mathbb{R}^2$.

4. Prendiamo adesso $U = \{(x,y) \in \mathbb{R}^2 \mid x \geq 0,\ y \geq 0\}$, cioè l'insieme dei punti del primo quadrante. In questo caso la somma di due vettori di U è ancora nel primo quadrante, l'elemento neutro $(0,0)$ ci sta, ma se consideriamo, ad esempio, $\lambda = -3$ e $u = (2,9)$, allora $\lambda u = (-6,-27)$ non appartiene ad U che quindi non è un sottospazio.

5. Si consideri $U = \{(x,y) \in \mathbb{R}^2 \mid x - 2y = 0\}$. Verifichiamo che U è un sottospazio vettoriale di $\mathbb{R}^2$. Siano $v_1 = (x_1, y_1)$ e $v_2 = (x_2, y_2)$ due elementi di U, tali che, cioè, $x_1 - 2y_1 = 0$ e $x_2 - 2y_2 = 0$. La somma dei due vettori: $v_1 + v_2 = (x_1 + x_2, y_1 + y_2)$ appartiene ad U, infatti $x_1 + x_2 - 2(y_1 + y_2) = (x_1 - 2y_1) + (x_2 - 2y_2) = 0 + 0 = 0$. Analogamente, per ogni numero reale λ, $\lambda v_1 = (\lambda x_1, \lambda y_1)$ è un elemento di U, dal momento che $\lambda x_1 - 2\lambda y_1 = \lambda(x_1 - 2y_1) = \lambda 0 = 0$. Usando la Proposizione 2.1.2 concludiamo che U è un sottospazio vettoriale di $\mathbb{R}^2$.

6. Sia $U = \{(x,y) \in \mathbb{R}^2 \mid yx = 0\}$ cioè l'insieme delle coppie (x,y) di numeri reali con almeno una componente uguale a 0. Osserviamo che l'insieme U non è chiuso rispetto alla somma, infatti i vettori $(1,0)$ e $(0,1)$ appartengono ad U ma non la loro somma: $(1,1)$. Dunque U non è un sottospazio vettoriale di $\mathbb{R}^2$.

7. Si prenda $U = \{(x, y) \in \mathbb{R}^2 \mid y + 3x - 1 = 0\}$. Allora U non è un sottospazio vettoriale di $\mathbb{R}^2$ dal momento che il vettore nullo di $\mathbb{R}^2$, $(0, 0)$, non appartiene ad U.

Osservazione sugli esempi fatti
Per stabilire se un sottoinsieme S di uno spazio vettoriale V è un sottospazio vettoriale di V possono essere utilizzati due approcci:

1°. Se pensiamo che l'insieme considerato non sia un sottospazio vettoriale basterà portare un esempio di vettori che non soddisfino le condizioni di 2.1.2 come fatto negli esempi 2, 3, 4, 6, 7.

2°. Se invece pensiamo che l'insieme considerato sia un sottospazio vettoriale allora si procederà come nell'esempio 5 verificando la validità delle condizioni di 2.1.2.

Nell'incertezza, ovvero quando non abbiamo a priori un'idea sul fatto che l'insieme in oggetto sia o meno un sottospazio vettoriale, si procede sempre come in 2°. Se l'insieme non è un sottospazio si arriva ad un ostacolo il che consente di trovare un controesempio. Ad esempio, consideriamo l'insieme dell'esempio 3. Procedendo come in 5, una delle condizioni da verificare è che per ogni $v = (x, y)$ elemento di $U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ e $\forall \lambda \in \mathbb{R}$ si ha $\lambda v \in U$. Ora da $v \in U$ sappiamo che $x^2 + y^2 \leq 1$ quindi $\lambda v = (\lambda x, \lambda y)$ e $(\lambda x)^2 + (\lambda y)^2 = \lambda^2(x^2 + y^2) \leq \lambda^2$ ma $\lambda v \in U$ se e solo se $\lambda^2(x^2 + y^2) \leq 1$. Ora è evidente che se v è tale che $x^2 + y^2 = 1$, ad esempio $v = (1, 0)$, per λ "grandi" $\lambda v \notin U$. Ad esempio $v = (1, 0) \in U$, ma preso $\lambda = 100$, $100v = (100, 0) \notin U$ perché $100^2 > 1$.

2.2 Generatori

Sia V un $\mathbb{R}$-spazio vettoriale e siano $v_1, v_2, \ldots, v_n$ vettori di V e $\lambda_1, \lambda_2, \ldots, \lambda_n$ numeri reali. Allora possiamo considerare per ogni indice $i = 1, 2, \ldots, n$ il vettore $\lambda_i v_i \in V$ e poi fare la somma in V dei vettori ottenuti:

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

(si noti che, per la proprietà commutativa della somma, l'ordine in cui vengono sommati gli elementi non altera il risultato finale).

Definizione 2.2.1 *Dato un $\mathbb{R}$-spazio vettoriale V e dato un insieme di vettori $v_1, v_2, \ldots, v_n$ di V, si dirà che un vettore $v \in V$ è loro* **combinazione**

lineare *se esistono dei numeri reali* $\lambda_1, \lambda_2, \dots, \lambda_n$ *tali che*

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n.$$

Gli scalari $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ si dicono coefficienti della combinazione lineare.

Esempio 2.2.2 Si consideri lo spazio vettoriale $\mathbb{R}^2$.

1. Il vettore $(\sqrt{5}, \sqrt{5})$ è combinazione lineare dei vettori $(1, -1), (0, \sqrt{5})$ in quanto $(\sqrt{5}, \sqrt{5}) = \sqrt{5}(1, -1) + 2(0, \sqrt{5})$;
2. una generica combinazione lineare dei vettori $(1, 0)$, $(0, 1)$ è $a(1, 0) + b(0, 1) = (a, b)$ con $a, b \in \mathbb{R}$. Quindi ogni vettore di $\mathbb{R}^2$ è combinazione lineare dei vettori $(1, 0), (0, 1)$.

In ogni spazio vettoriale il vettore nullo è combinazione lineare di qualunque insieme di vettori. Infatti se $v_1, \dots, v_n$ è un insieme di vettori in V spazio vettoriale, allora il vettore nullo $\mathbf{0}_V$ è uguale alla combinazione lineare nulla $\mathbf{0}_V = 0v_1 + 0v_2 + \cdots + 0v_n$.

Che cosa succederebbe se tutti i vettori di V si potessero scrivere come combinazione lineare di un numero finito di vettori come nell'esempio 2 di 2.2.2?

Definizione 2.2.3 *Un* $\mathbb{R}$*-spazio vettoriale* V *si dice* **finitamente generato** *se esiste un insieme finito di vettori* $v_1, v_2, \dots, v_n$ *tali che ogni vettore di* V *sia loro combinazione lineare. I vettori* $v_1, v_2, \dots, v_n$ *si dicono allora* **generatori** *di* V*.*

Esempio 2.2.4 Prendiamo la nostra vecchia conoscenza $\mathbb{R}^3$ e tre suoi vettori $u_1 = (1, 1, -1)$, $u_2 = (0, 1, 2)$, $u_3 = (0, -2, 3)$. Vogliamo dimostrare che $\mathbb{R}^3$ è finitamente generato mostrando che $\{u_1, u_2, u_3\}$ è un insieme di generatori di $\mathbb{R}^3$. Dobbiamo verificare che **ogni** vettore di $\mathbb{R}^3$ è combinazione lineare di u_1, u_2, u_3 . Prendiamo un vettore qualsiasi di $\mathbb{R}^3$: (α, β, γ) con $\alpha, \beta, \gamma \in \mathbb{R}$. Dobbiamo provare che esistono ben definiti $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tali che

$$(\alpha, \beta, \gamma) = \lambda_1 u_1 +_{\mathbb{R}^3} \lambda_2 u_2 +_{\mathbb{R}^3} \lambda_3 u_3.$$

Esplicitiamo tale somma usando la somma ed il prodotto per scalari che abbiamo definito in $\mathbb{R}^3$:

$$(\alpha, \beta, \gamma) \quad = \quad \lambda_1(1, 1, -1) +_{\mathbb{R}^3} \lambda_2(0, 1, 2) +_{\mathbb{R}^3} \lambda_3(0, -2, 3) =$$

$$\begin{aligned} &= (\lambda_1, \lambda_1, -\lambda_1) +_{\mathbb{R}^3} (0, \lambda_2, 2\lambda_2) +_{\mathbb{R}^3} (0, -2\lambda_3, 3\lambda_3) = \\ &= (\lambda_1, \lambda_1 + \lambda_2 - 2\lambda_3, -\lambda_1 + 2\lambda_2 + 3\lambda_3). \end{aligned}$$

Per calcolare i λ_i richiesti dobbiamo risolvere il sistema

$$\begin{cases} \alpha = \lambda_1 \\ \beta = \lambda_1 + \lambda_2 - 2\lambda_3 \\ \gamma = -\lambda_1 + 2\lambda_2 + 3\lambda_3. \end{cases}$$

Il sistema ottenuto ha soluzione: $\lambda_1 = \alpha$, $\lambda_2 = \frac{2}{7}\gamma - \frac{1}{7}\alpha + \frac{3}{7}\beta$ e $\lambda_3 = \frac{1}{7}(\gamma + 3\alpha - 2\beta)$, quindi i vettori u_1, u_2, u_3 sono dei generatori di $\mathbb{R}^3$. Si noti che anche l'insieme $\{(2,1,0),(0,1,0),(0,0,1)\}$ è un insieme di generatori di $\mathbb{R}^3$ come pure $\{(1,1,0),(0,1,1),(1,0,1),(2,1,1)\}$. Quindi gli insiemi di generatori non sono unici.

Esempio 2.2.5 Non tutti gli spazi vettoriali sono finitamente generati. Prendiamo ad esempio lo spazio vettoriale dei polinomi $\mathbb{R}[X]$ nella indeterminata X. Se fosse finitamente generato esisterebbe un numero finito di polinomi $P_1(X), \ldots, P_n(X)$ che generano $\mathbb{R}[X]$, quindi ogni polinomio a coefficienti reali si scriverebbe come loro combinazione lineare. Sia g_i con $i = 1, \ldots, n$, il grado del polinomio $P_i(x)$ e sia g il massimo tra questi gradi. Allora ogni combinazione lineare di $P_1(X), \ldots, P_n(X)$ avrà al più grado g e non potrà mai avere grado superiore. In particolare un polinomio di grado $g + 1$ non si potrà mai scrivere come combinazione lineare di $P_1(X), \ldots, P_n(X)$. Dunque i polinomi $P_1(X), \ldots, P_n(X)$ non sono dei generatori di $\mathbb{R}[X]$ per nessun n.

Analogamente qualunque spazio vettoriale di cui $\mathbb{R}[X]$ sia sottospazio non sarà finitamente generato.

Osservazione 2.2.6 Sia $\mathbf{u} = \{u_1, u_2, \ldots, u_s\}$ un insieme di generatori di un $\mathbb{R}$-spazio vettoriale V. Allora ogni vettore v di V si scrive come loro combinazione lineare:

$$v = \lambda_1 u_1 +_V \lambda_2 u_2 +_V \ldots +_V \lambda_s u_s.$$

Tale scrittura è unica? In generale la risposta è no. Ad esempio l'insieme $\{(0),(1)\}$ è chiaramente un insieme di generatori dello spazio vettoriale reale $\mathbb{R}$ (perché ogni $(\alpha) \in \mathbb{R}$ è $(\alpha) = \alpha(1) + (0)$, ma la scrittura non è unica infatti $(0) = 0(1) + 0(0) = 0(1) + 1(0)$. Come ulteriore esempio si consideri l'insieme ordinato $\mathcal{B} = \{(1,1,0),(0,1,1),(1,0,1),(2,1,1)\}$ di generatori di $\mathbb{R}^3$. Il vettore $(0,1,1)$ si può scrivere come $0(1,1,0) + 1(0,1,1) + 0(1,0,1) + 0(2,1,1)$, ma pure come $1(1,1,0) + 1(0,1,1) + 1(1,0,1) - 1(2,1,1)$.

Osservazione 2.2.7 Osserviamo che, dato un $\mathbb{R}$-spazio vettoriale V generato dall'insieme di vettori $\{u_1, u_2, \dots, u_s\}$ e preso qualsiasi vettore $u \in V$, l'insieme di vettori $\{u_1, u_2, \dots, u_s, u\}$ è pure un insieme di generatori di V. Ovviamente si possono continuare ad aggiungere vettori ottenendo sempre un insieme di generatori.

Definizione 2.2.8 *Siano V un $\mathbb{R}$-spazio vettoriale e $\{u_1, u_2, \dots, u_k\}$ un insieme di vettori di V. Il più piccolo sottospazio di V contenente $u_1, u_2, \dots, u_k$ si dirà il sottospazio da essi generato e si indicherà con $\langle u_1, u_2, \dots, u_k\rangle$. Tale sottospazio è costituito da tutte le combinazioni lineari di $u_1, u_2, \dots, u_k$:*

$$\langle u_1, u_2, \dots, u_k\rangle = \left\{ \sum_{i=1}^{k} \alpha_i u_i = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k \mid \alpha_i \in \mathbb{R}, i = 1, \dots k \right\}$$

Il contenuto di questa definizione va giustificato. Prima di tutto verificheremo che l'insieme formato da tutte le combinazioni lineari di $u_1, u_2, \dots, u_k$ è un sottospazio vettoriale di V; poi vedremo che esso è contenuto in ogni sottospazio contenente $u_1, u_2, \dots, u_k$, e che quindi è il più piccolo sottospazio soddisfacente questa proprietà. Siano $\lambda \in \mathbb{R}$ e $v = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k$ una combinazione lineare di $u_1, u_2, \dots, u_k$. Allora $\lambda v = \lambda(\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k) = \lambda\alpha_1 u_1 + \lambda\alpha_2 u_2 + \ldots + \lambda\alpha_k u_k$ è pure una combinazione lineare di $u_1, u_2, \dots, u_k$ e quindi un elemento del nostro insieme. Se poi abbiamo due combinazioni lineari $v = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k$ e $w = \beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_k u_k$, allora la loro somma $v + w = (\alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_k u_k) + (\beta_1 u_1 + \beta_2 u_2 + \ldots + \beta_k u_k) = (\alpha_1 + \beta_1)u_1 + (\alpha_2 + \beta_2)u_2 + \ldots + (\alpha_k + \beta_k)u_k$ è ancora una combinazione lineare dei vettori $u_1, u_2, \dots, u_k$, quindi appartiene al nostro insieme che è pertanto un sottospazio vettoriale di V. D'altro canto se consideriamo un sottospazio $T \leq V$ contenente i vettori $u_1, u_2, \dots, u_k$, per definizione di sottospazio, T contiene ogni loro combinazione lineare e quindi tutto $\langle u_1, u_2, \dots, u_k\rangle$.

Osserviamo che, vista la definizione 2.2.8, è chiaro che uno spazio vettoriale V è finitamente generato se e solo se esistono $v_1, \dots, v_n$ vettori di V tali che $V = \langle v_1, \dots, v_n\rangle$.

Se invece di partire da un numero finito di vettori $\{u_1, u_2, \dots, u_k\}$ di uno spazio vettoriale V partissimo da un generico sottoinsieme $S \subset V$, potremmo chiederci qual è il più piccolo sottospazio vettoriale di V che contiene S?

Definizione 2.2.9 *Dato un qualsiasi sottoinsieme S di uno spazio vettoriale reale V, indichiamo con $\langle S\rangle$ il più piccolo sottospazio vettoriale di V contenente S. Allora $\langle S\rangle$ è l'insieme di tutte le combinazioni lineari di elementi di S a coefficienti in $\mathbb{R}$.*

Esempio 2.2.10 Sia V un $\mathbb{R}$-spazio vettoriale, allora $\emptyset \subset V$ e $\langle\emptyset\rangle = \{\mathbf{0}_V\}$. Ma anche $\{\mathbf{0}_V\} \subset V$ e $\langle\mathbf{0}_V\rangle = \{\mathbf{0}_V\}$.

Esempio 2.2.11 Dato un $\mathbb{R}$-spazio vettoriale V, lo spazio vettoriale generato da un vettore non nullo v è l'insieme $\langle v\rangle = \{\alpha v \mid \alpha \in \mathbb{R}\}$ di tutti i multipli scalari di v.

2.3 Operazioni tra sottospazi

Dato un insieme X è possibile definire delle operazioni sui suoi sottoinsiemi. In particolare se P_1, P_2 sono sottoinsiemi di X l'intersezione $P_1 \cap P_2$ è il più grande sottoinsieme di X contenuto in entrambi, e l'unione $P_1 \cup P_2$ è il più piccolo sottoinsieme di X che contiene entrambi. Possiamo tentare di definire le stesse operazioni sostituendo il termine "insieme" con "spazio vettoriale". Il progetto ha facilmente successo nel caso dell'intersezione:

Proposizione 2.3.1 *In un $\mathbb{R}$-spazio vettoriale V l'intersezione (insiemistica) $S_1 \cap S_2$ di due sottospazi S_1 e S_2 è un sottospazio vettoriale di V.*

Dimostrazione. Si deve mostrare che $S_1 \cap S_2$ è un sottospazio di V: per questo useremo la caratterizzazione 2.1.2. Siano quindi $v \in S_1 \cap S_2$ e $\lambda \in \mathbb{R}$. In particolare $v \in S_1$, che è un sottospazio, quindi $\lambda v \in S_1$; analogamente $v \in S_2$, che è un sottospazio, quindi $\lambda v \in S_2$ da cui λv appartiene sia a S_1 che a S_2 quindi appartiene alla loro intersezione. Siano $v_1, v_2 \in S_1 \cap S_2$, in particolare $v_1, v_2 \in S_1$ che è un sottospazio e quindi $v_1 +_V v_2 \in S_1$, analogamente $v_1, v_2 \in S_2$ che è un sottospazio e quindi $v_1 +_V v_2 \in S_2$, si conclude che $v_1 +_V v_2 \in S_1 \cap S_2$. **C.V.D.**

Nel caso dell'unione vi sono dei problemi, infatti l'unione insiemistica di due sottospazi di uno spazio vettoriale V in generale non è un sottospazio di V. Prendiamo ad esempio in $\mathbb{R}^2$ i due sottospazi $W_1 = \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$ e $W_2 = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}$. Allora la loro unione insiemistica è l'insieme dei vettori in cui o le due coordinate sono uguali o sono opposte:

$$W_1 \cup W_2 = \{(x, y) \in \mathbb{R}^2 \mid (x - y)(x + y) = 0\}$$

(si ricordi che il prodotto di due numeri reali è zero se almeno uno dei due numeri è zero). L'insieme $W_1 \cup W_2$ non è un sottospazio vettoriale di $\mathbb{R}^2$ perché, pur essendo stabile rispetto alla moltiplicazione per uno scalare, non è chiuso rispetto alla somma: $(2,2)+(3,-3) = (5,-1)$ e le coordinate di quest'ultima coppia non sono né uguali né opposte! Quindi l'unione insiemistica di due sottospazi vettoriali di uno spazio V non ha necessariamente la struttura di spazio vettoriale. Più in generale vale:

Proposizione 2.3.2 *Siano V un $\mathbb{R}$-spazio vettoriale e W_1, W_2 due suoi sottospazi, allora $W_1 \cup W_2$ è un sottospazio di V se e solo se $W_1 \subseteq W_2$ o $W_2 \subseteq W_1$.*

Dimostrazione. "$\Leftarrow$" Se $W_1 \subseteq W_2$ (risp. se $W_2 \subseteq W_1$), allora $W_1 \cup W_2 = W_2$ (risp. $W_1 \cup W_2 = W_1$) è sottospazio vettoriale per ipotesi.

"$\Rightarrow$" Per provare questa implicazione proviamo che se $W_1 \nsubseteq W_2$ e $W_2 \nsubseteq W_1$ allora $W_1 \cup W_2$ non è sottospazio. Essendo $W_1 \nsubseteq W_2$ esiste $v_1 \in W_1 \setminus W_2$, analogamente essendo $W_2 \nsubseteq W_1$ esiste $v_2 \in W_2 \setminus W_1$. Se $W_1 \cup W_2$ fosse sottospazio allora $v = v_1 + v_2$ dovrebbe essere un elemento di $W_1 \cup W_2$ in quanto somma di un elemento di W_1 e di uno di W_2. Se v fosse in W_1 allora anche $v_2 = v - v_1$ apparterrebbe a W_1, ma avevamo scelto $v_2 \in W_2 \setminus W_1$. Analogamente se v fosse in W_2 allora anche $v_1 = v - v_2$ apparterrebbe a W_2, ma avevamo scelto $v_1 \in W_1 \setminus W_2$. Quindi $v \notin W_1 \cup W_2$. **C.V.D.**

Abbiamo dunque stabilito che il più piccolo INSIEME di vettori di V che contiene i due sottospazi non è in generale un sottospazio vettoriale di V. Cerchiamo allora il più piccolo SOTTOSPAZIO di V che contenga W_1 e W_2. Memori della definizione 2.2.9 tale sottospazio sarà il sottospazio vettoriale generato da W_1 e da W_2. Consideriamo la seguente costruzione:

Definizione 2.3.3 *Sia V un $\mathbb{R}$-spazio vettoriale e siano S_1, S_2 due suoi sottospazi. Si definisce somma di S_1 ed S_2 l'insieme:*

$$S_1 + S_2 = \{v_1 +_V v_2 \mid v_1 \in S_1, v_2 \in S_2\}.$$

In altre parole $S_1 + S_2$ è l'insieme dei vettori di V che sono somma di un vettore di S_1 e di un vettore di S_2. Si noti che l'insieme $S_1 + S_2$ contiene S_1 poiché ogni vettore v di S_1 si scrive come $v +_V \mathbf{0}_V \in S_1 + S_2$ (in quanto l'elemento neutro di V appartiene a S_2 che è un sottospazio) e, per lo stesso motivo, contiene S_2.

Proposizione 2.3.4 *In un $\mathbb{R}$-spazio vettoriale V, dati due sottospazi S_1 e S_2, l'insieme $S_1 + S_2$ è un sottospazio vettoriale di V ed è il più piccolo sottospazio di V che contiene sia S_1 che S_2.*

Dimostrazione. Mostriamo innanzitutto che $S_1 + S_2$ è un sottospazio di V. Usiamo il criterio 2.1.2. Prendiamo un elemento $v_1 + v_2 \in S_1 + S_2$ (con $v_1 \in S_1$ e $v_2 \in S_2$) e un numero reale $\lambda \in \mathbb{R}$. Ci chiediamo: $\lambda(v_1+v_2)$ è ancora un elemento di $S_1 + S_2$? In effetti, per la proprietà distributiva del prodotto per scalari rispetto alla somma, si ha $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$, ma $\lambda v_1 \in S_1$ e $\lambda v_2 \in S_2$ perché S_1 e S_2 sono sottospazi. Dunque $\lambda(v_1 + v_2) \in S_1 + S_2$. Ancora, siano v e w elementi di $S_1 + S_2$, cioè $v = v_1 + v_2$ e $w = w_1 + w_2$ con $v_1, w_1 \in S_1$ e $v_2, w_2 \in S_2$: allora $v + w$ è un elemento di $S_1 + S_2$? Si ha: $v + w = (v_1 + v_2) + (w_1 + w_2)$ che, per la commutatività e la associatività della somma, è uguale a $(v_1 + w_1) + (v_2 + w_2)$; ora $v_1 + w_1 \in S_1$ perché S_1 è un sottospazio, analogamente $v_2 + w_2 \in S_2$ e possiamo concludere che $v + w \in S_1 + S_2$. Osserviamo che $S_1 + S_2$ è il più piccolo sottospazio di V che abbia la proprietà di contenere sia S_1 che S_2: infatti, se un sottospazio $T \leq V$ contiene sia S_1 che S_2 allora dovrà necessariamente contenere, per definizione di sottospazio, tutte le somme $v_1 + v_2$, al variare di $v_1 \in S_1$ e $v_2 \in S_2$, cioè $S_1 + S_2$. **C.V.D.**

2.4 Esercizi svolti

Esercizio 2.4.1 Stabilire se i seguenti insiemi di vettori generano $\mathbb{R}^3$:

(i) $(0, 0, 1)$, $(2, 1, 0)$, $(1, 1, 1)$;

(ii) $(2, 3, 4)$, $(3, 2, 1)$;

(iii) $(1, 0, 0)$, $(1, 2, 0)$, $(0, 0, 2)$, $(1, 3, 4)$.

Svolgimento. L'esercizio consiste nel stabilire se ogni vettore di $\mathbb{R}^3$ si possa scrivere come combinazione lineare dei vettori di volta in volta indicati.

(i) Sia (α, β, γ) un generico elemento di $\mathbb{R}^3$. Ci chiediamo se esistono λ_1, λ_2, λ_3 in $\mathbb{R}$ tali che sia $(\alpha, \beta, \gamma) = \lambda_1(0, 0, 1) + \lambda_2(2, 1, 0) + \lambda_3(1, 1, 1)$, cioè $(\alpha, \beta, \gamma) = (2\lambda_2 + \lambda_3, \lambda_2 + \lambda_3, \lambda_1 + \lambda_3)$. Si tratta di risolvere il seguente sistema:

$$\begin{cases} \alpha = 2\lambda_2 + \lambda_3 \\ \beta = \lambda_2 + \lambda_3 \\ \gamma = \lambda_1 + \lambda_3. \end{cases}$$

Sottraendo la seconda equazione dalla prima otteniamo: $\lambda_2 = \alpha - \beta$; sottraendo la prima equazione da due volte la seconda otteniamo: $\lambda_3 = 2\beta - \alpha$. Infine, sostituendo nella terza equazione l'espressione ottenuta di λ_3, otteniamo: $\lambda_1 = \gamma - 2\beta + \alpha$. Dunque per ogni $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ abbiamo determinato i numeri reali λ_i che cercavamo. Pertanto l'insieme di vettori (i) genera tutto $\mathbb{R}^3$.

(ii) In questo esercizio l'insieme che ci viene proposto contiene solo due elementi. Anche in questo caso potremmo procedere come prima cercando di risolvere un sistema, ma questa volta il sistema avrà 3 equazioni e 2 incognite. Quindi è lecito aspettarsi che un tale sistema non abbia sempre soluzioni, cioè che vi siano dei vettori di $\mathbb{R}^3$ che non sono generati dall'insieme (ii). In questo caso per risolvere il quesito sarà sufficiente esibire un vettore di $\mathbb{R}^3$ che non è combinazione lineare di $(2, 3, 4)$ e $(3, 2, 1)$.

Consideriamo il vettore $(1, 0, 0)$ di $\mathbb{R}^3$. Ci chiediamo: è possibile scrivere $(1, 0, 0)$ come combinazione lineare di $(2, 3, 4)$, $(3, 2, 1)$? Se ciò fosse possibile esisterebbero λ_1 e λ_2 in $\mathbb{R}$ tali che $(1, 0, 0) = \lambda_1(2, 3, 4) + \lambda_2(3, 2, 1) = (2\lambda_1 + 3\lambda_2, 3\lambda_1 + 2\lambda_2, 4\lambda_1 + \lambda_2)$. Quindi il sistema

$$\begin{cases} 1 = 2\lambda_1 + 3\lambda_2 \\ 0 = 3\lambda_1 + 2\lambda_2 \\ 0 = 4\lambda_1 + \lambda_2. \end{cases}$$

dovrebbe ammettere soluzione, ma dalla seconda e dalla terza equazione ricaviamo $\lambda_1 = 0 = \lambda_2$ che non soddisfano la prima equazione. Il sistema non ha soluzioni cioè il vettore $(1, 0, 0)$ non è combinazione lineare dei vettori $(2, 3, 4)$ e $(3, 2, 1)$. Questo dimostra che $(2, 3, 4)$ e $(3, 2, 1)$ non generano $\mathbb{R}^3$.

(iii) Anche in questo caso potremmo precedere come in (i) e, preso $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, determinare λ_1, λ_2, λ_3, λ_4 in $\mathbb{R}$ tali che $(\alpha, \beta, \gamma) = \lambda_1(1, 0, 0) + \lambda_2(1, 2, 0) + \lambda_3(0, 0, 2) + \lambda_4(1, 3, 4)$.

Mostriamo invece un modo alternativo di procedere. Prima di tutto osserviamo un fatto di carattere generale: dato un insieme I di vettori di uno spazio vettoriale V e fissato un insieme S di generatori di V, se ogni elemento di S è combinazione lineare degli elementi di I, allora anche I è un insieme di generatori di V. Infatti ogni vettore di V si scrive come combinazione lineare dei vettori di S che a loro volta sono combinazioni lineari dei vettori di I e quindi ogni vettore di V è combinazione lineare dei vettori di I. Osserviamo inoltre che i vettori $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ generano $\mathbb{R}^3$, infatti ogni elemento (α,β,γ) è loro combinazione lineare:

$$(\alpha,\beta,\gamma) = \alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1).$$

Notiamo, dunque, che il vettore $(1,0,0)$ appartiene all'insieme dato e che si ha: $(0,1,0) = \frac{1}{2}((1,2,0) - (1,0,0))$; $(0,0,1) = \frac{1}{2}(0,0,2)$. Concludiamo, per quanto appena osservato, che l'insieme di vettori (iii) genera $\mathbb{R}^3$. Esplicitamente per un vettore qualsiasi (α,β,γ) di $\mathbb{R}^3$ si ha $(\alpha,\beta,\gamma) = \alpha(1,0,0) + \beta(0,1,0) + \gamma(0,0,1) = \alpha(1,0,0) + \beta(\frac{1}{2}((1,2,0) - (1,0,0))) + \gamma(\frac{1}{2}(0,0,2)) = (\alpha - \frac{\beta}{2})(1,0,0) + \frac{\beta}{2}(1,2,0) + \frac{\gamma}{2}(0,0,2)$.

Esercizio 2.4.2 Determinare due insiemi distinti di generatori dello spazio vettoriale $\mathbb{R}^{\leq 2}[x]$ dei polinomi a coefficienti reali nella variabile x di grado ≤ 2.

Svolgimento. L'insieme $\{1, x, x^2\}$ è certamente un insieme di generatori di $\mathbb{R}^{\leq 2}[x]$. Infatti ogni elemento di $\mathbb{R}^{\leq 2}[x]$ è un polinomio della forma $a+bx+cx^2$ con $a, b, c \in \mathbb{R}$ quindi $a + bx + cx^2 = a \cdot 1 + b \cdot x + c \cdot x^2$.

È molto facile, a questo punto, costruire un altro insieme di generatori: basterà aggiungere all'insieme già individuato un qualsiasi altro polinomio di $\mathbb{R}^{\leq 2}[x]$. Così, ad esempio, l'insieme $\{1, x, x^2, 4x + 69x^2\}$ genera $\mathbb{R}^{\leq 2}[x]$. Ma pure $\{1, x, 4x + 69x^2\}$ è un insieme di generatori, mentre $\{x, x^2, 4x + 69x^2\}$ non lo è (verificarlo per esercizio).

Esercizio 2.4.3 Mostrare che l'insieme dei polinomi $3 + x$, x^2, $1 + x^2 + x^3$ non è un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$.

Svolgimento. Prendiamo un monomio di grado 0, ad esempio 1, e vediamo se è possibile scriverlo come combinazione lineare dei polinomi $3 + x$, x^2, $1 + x^2 + x^3$, cerchiamo cioè α, β, γ tali che sia $1 = \alpha(3 + x) + \beta x^2 + \gamma(1 + x^2 + x^3) = 3\alpha + \gamma + \alpha x + (\beta + \gamma)x^2 + \gamma x^3$. Ricordiamo che due polinomi sono uguali se i coefficienti dei termini dello stesso grado dei due polinomi sono ordinatamente uguali. Dobbiamo pertanto risolvere il sistema

$$\begin{cases} 1 = 3\alpha + \gamma \\ 0 = \alpha \\ 0 = \beta + \gamma \\ 0 = \gamma. \end{cases}$$

Si vede immediatamente che il sistema trovato non ha soluzioni, pertanto i polinomi $3 + x$, x^2, $1 + x^2 + x^3$ non individuano un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$.

Esercizio 2.4.4 Verificare che l'insieme delle matrici $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ genera lo spazio vettoriale $\mathcal{M}_2(\mathbb{R})$.

Svolgimento. Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una qualunque matrice di $\mathcal{M}_2(\mathbb{R})$. Allora $A = ae_{11} + be_{12} + ce_{21} + de_{22}$ (notiamo inoltre che tale scrittura è unica). Questo prova che $\{e_{11}, e_{12}, e_{21}, e_{22}\}$ è un insieme di generatori di $\mathcal{M}_2(\mathbb{R})$. Osserviamo che i coefficienti della combinazione lineare trovata coincidono con le entrate di A. Un altro insieme di generatori di $\mathcal{M}_2(\mathbb{R})$ è l'insieme costituito dalle matrici $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $T = \begin{pmatrix} 3 & 0 \\ 2 & 1 \end{pmatrix}$.

Esercizio 2.4.5 L'insieme dei polinomi di grado 2 a coefficienti reali nella variabile x è un sottospazio vettoriale di $\mathbb{R}[x]$?

Svolgimento. Condizione necessaria affinché un sottoinsieme S di uno spazio vettoriale V sia un sottospazio di V è che S contenga il vettore nullo $\mathbf{0}_V$ di V. Il vettore nullo nello spazio vettoriale $\mathbb{R}[x]$ è il polinomio identicamente nullo, cioè il polinomio di grado 0 con termine noto uguale a 0, pertanto esso non è contenuto nell'insieme dei polinomi di grado 2. Concludiamo che l'insieme dei polinomi di grado 2 non è un sottospazio vettoriale di $\mathbb{R}[x]$.

Esercizio 2.4.6 Stabilire quali tra i seguenti sottoinsiemi sono sottospazi vettoriali di $\mathbb{R}^3$:

(i) $A = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}$;

(ii) $B = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$;

(iii) $C = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y = 0\}$.

Svolgimento.

(i) Come nell'esercizio precedente, osserviamo immediatamente che l'insieme A non contiene il vettore $\mathbf{0}_{\mathbb{R}^3} = (0,0,0)$ pertanto A non è un sottospazio vettoriale di $\mathbb{R}^3$.

Gli insiemi B e C contengono $(0,0,0)$ ma questo non è sufficiente a provare che essi siano sottospazi vettoriali di $\mathbb{R}^3$. Dobbiamo stabilire se B e C sono chiusi rispetto alle operazioni di $\mathbb{R}^3$, cioè se presi due vettori v, w in B (risp. C) la loro somma appartiene ancora a B (risp. C) e se preso un qualunque vettore v in B (risp. C) e un qualunque numero reale λ il prodotto λv appartiene a B (risp. C).

(ii) Siano $v = (x, y, z)$ e $w = (\tilde{x}, \tilde{y}, \tilde{z})$ elementi di B, cioè: $x + y + z = 0$ e $\tilde{x} + \tilde{y} + \tilde{z} = 0$. Allora il vettore $v + w = (x + \tilde{x}, y + \tilde{y}, z + \tilde{z})$ appartiene a B dal momento che le sue componenti soddisfano l'equazione di B: $(x + \tilde{x}) + (y + \tilde{y}) + (z + \tilde{z}) = (x + y + z) + (\tilde{x} + \tilde{y} + \tilde{z}) = 0 + 0 = 0$. Analogamente, per ogni $\lambda \in \mathbb{R}$, $\lambda v = (\lambda x, \lambda y, \lambda z)$ appartiene a B dal momento che $\lambda x + \lambda y + \lambda z = \lambda(x + y + z) = \lambda 0 = 0$. Quindi B è sottospazio vettoriale di $\mathbb{R}^3$.

(iii) Consideriamo ora l'insieme C: i vettori $v = (1, -1, 0)$ e $w = (-1, -1, 0)$ appartengono a C ma la loro somma $v + w = (0, -2, 0)$ non appartiene a C (perché $0^2 - 2 \neq 0$). Pertanto C non è chiuso rispetto alla somma e quindi non è un sottospazio vettoriale di $\mathbb{R}^3$.

Quest'ultimo esercizio ci fornisce un'indicazione che potrà essere utile: se un sottoinsieme di uno spazio vettoriale è descritto da equazioni, tale sottoinsieme difficilmente sarà un sottospazio vettoriale se le equazioni coinvolte non sono lineari nelle incognite oppure se appare un termine noto. Per ora prenderemo questa osservazione solo come un'indicazione di massima.

Esercizio 2.4.7 Verificare che i seguenti insiemi di matrici sono sottospazi vettoriali di $\mathcal{M}_3(\mathbb{R})$ e per ciascuno di essi esibire un insieme di generatori.

(i) $A = \left\{ \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}$;

(ii) $B = \left\{ \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mid \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\}$;

(iii) $C = \left\{ \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \mid a, b, c, d, e, f \in \mathbb{R} \right\}$.

Svolgimento. Come nell'esercizio precedente, per verificare che A, B e C sono sottospazi vettoriali di $\mathcal{M}_3(\mathbb{R})$ basta utilizzare la proposizione 2.1.2. Effettuiamo questa verifica soltanto per l'insieme A e lasciamo allo studente la verifica per gli insiemi B e C.

(i) L'insieme A è l'insieme delle matrici quadrate $M = (m_{ij})$ di ordine 3 $(i, j \in \{1, 2, 3\})$ ad entrate reali, triangolari strettamente superiori, cioè delle matrici quadrate di ordine 3 per cui $m_{ij} = 0$ se $i \geq j$ (i.e. le matrici i cui elementi diagonali sono nulli e i cui elementi al di sotto della diagonale principale sono anch'essi nulli). Siano dunque $M_1 = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$ e $M_2 = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 0 \end{pmatrix}$ due elementi di A, con $a, b, c, \alpha, \beta, \gamma \in \mathbb{R}$. Allora la matrice $M_1 + M_2 = \begin{pmatrix} 0 & a+\alpha & b+\beta \\ 0 & 0 & c+\gamma \\ 0 & 0 & 0 \end{pmatrix}$ è evidentemente una matrice triangolare strettamente superiore e quindi appartiene all'insieme A. Analogamente, per ogni $\lambda \in \mathbb{R}$, $\lambda M_1 = \begin{pmatrix} 0 & \lambda a & \lambda b \\ 0 & 0 & \lambda c \\ 0 & 0 & 0 \end{pmatrix}$ appartiene all'insieme A. Pertanto A è chiuso rispetto alla somma e al prodotto per scalari ed quindi è un sottospazio vettoriale di $\mathcal{M}_3(\mathbb{R})$. Per determinare un insieme di generatori di A osserviamo che un suo generico elemento $M = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$, con $a, b, c \in \mathbb{R}$ si può sempre scrivere come:

$$M = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} = a \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

e quindi l'insieme delle matrici $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ genera il sottospazio A.

(ii) Procedendo nello stesso modo otteniamo che $\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$ è un insieme di generatori del sottospazio B.

(iii) Infine, le matrici $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ generano C.

C è detto l'insieme delle *matrici simmetriche* di ordine 3.

2.5 Esercizi proposti

Esercizio 2.5.1 Costruire un esempio di un sottoinsieme di $\mathcal{M}_2(\mathbb{R})$ costituito da infiniti elementi, che non sia un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$.

Esercizio 2.5.2 Si considerino i seguenti sottoinsiemi S e T di $\mathbb{R}^4$:

$$S = \{(a, b, 0, 0) \mid a, b \in \mathbb{R}\}, \qquad T = \{(c, 0, 0, d) \mid c, d \in \mathbb{R}\}.$$

1. Dimostrare che S e T sono sottospazi vettoriali di $\mathbb{R}^4$.
2. Determinare $S \cap T$ e $S + T$.
3. Mostrare che ogni vettore di $S+T$ si scrive in modi diversi come somma di vettori di S e T.

Esercizio 2.5.3 Determinare i valori di $k \in \mathbb{R}$ tali che l'insieme

$$S_k = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + 2y - kz + 8t = k\}$$

sia un sottospazio vettoriale di $\mathbb{R}^4$. Per i valori di k trovati determinare un insieme di generatori di S_k e, se possibile, esibire un vettore di $\mathbb{R}^4$ che non sia una combinazione lineare di questi.

Esercizio 2.5.4 Sia S l'insieme dei polinomi di grado 3 a coefficienti reali nella variabile x.

1. S è un sottospazio vettoriale di $\mathbb{R}[x]$?

2. Determinare il più piccolo sottospazio vettoriale T di $\mathbb{R}[x]$ contenente S.

Esercizio 2.5.5 Sia $S = \{(x, y) \in \mathbb{R}^2 \mid 2x^2 - 7xy + 3y^2 = 0\}$.

1. Mostrare che S non è un sottospazio vettoriale di $\mathbb{R}^2$.

2. Scrivere, se possibile, S come unione di sottospazi vettoriali di $\mathbb{R}^2$.

3. Determinare il più piccolo sottospazio vettoriale di $\mathbb{R}^2$ contenente S.

Lezione 3

Basi e dimensione

3.1 Dipendenza e indipendenza lineare

In 2.2.7 abbiamo visto che, dato un $\mathbb{R}$-spazio vettoriale V generato dai vettori u_1, u_2, …, u_s (quindi $V = \langle u_1, u_2, \ldots, u_s \rangle$) si possono aggiungere all'insieme $\{u_1, u_2, \ldots, u_s\}$ altri vettori e ottenere ancora un insieme di generatori. Inoltre abbiamo visto che ogni vettore $v \in V$ si può scrivere come combinazione lineare di u_1, u_2, …, u_s tramite $\lambda_1, \lambda_2, \ldots, \lambda_s \in \mathbb{R}$: $v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_s u_s$, ma che tali λ_i non sono in generale unici. In questo paragrafo cercheremo di caratterizzare gli insiemi di vettori I per i quali valga il fatto che se un vettore si scrive come combinazione lineare dei vettori di I allora tale scrittura è unica. Se questi vettori sono anche dei generatori dello spazio vettoriale V potremo scrivere ogni vettore di V in maniera unica come loro combinazione lineare. Vedremo ora come per testare l'unicità della scrittura basti testarla per il vettore nullo $\mathbf{0}_V$.

Definizione 3.1.1 *In un $\mathbb{R}$-spazio vettoriale V i vettori $u_1, u_2, \ldots, u_k$, $k \in \mathbb{N}$, si dicono* **linearmente indipendenti** *se il solo modo di scrivere il vettore nullo $\mathbf{0}_V$ come loro combinazione lineare è con tutti i coefficienti nulli, i.e.*

$$\mathbf{0}_V = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k \Leftrightarrow \lambda_1 = \lambda_2 = \ldots = \lambda_k = 0.$$

Se, al contrario, il vettore nullo si può scrivere in modi diversi come combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$, allora diremo che i vettori $u_1, u_2, \ldots, u_k$ sono **linearmente dipendenti**.

Esempi 3.1.2 1. In $\mathbb{R}^3$ i due vettori $(1, 0, 1), (2, 1, 1)$ sono linearmente indipendenti: infatti se dovessimo scrivere il vettore nullo di $\mathbb{R}^3$ come loro combinazione lineare avremmo $(0,0,0) = \lambda_1(1,0,1)+\lambda_2(2,1,1)$ da cui otterremmo che $(0,0,0) = (\lambda_1 + 2\lambda_2, \lambda_2, \lambda_1 + \lambda_2)$ cioè $\lambda_1 = \lambda_2 = 0$.

2. Consideriamo ora in $\mathbb{R}^2$ l'insieme $I = \{(1,1),(2,1),(1,-1)\}$. Il vettore nullo si scrive in modi diversi come combinazione lineare dei vettori di I:

$$3(1,1) - 2(2,1) + (1,-1) = 0(1,1) + 0(2,1) + 0(1,-1) = (0,0),$$

quindi i vettori di I sono linearmente dipendenti.

3. In $\mathbb{R}^n$ consideriamo un insieme di vettori con la seguente proprietà: ciascun vettore ha una componente diversa da zero, diciamo la i-esima, e i rimanenti vettori hanno invece entrata i-esima nulla. Tali vettori sono linearmente indipendenti. Ad esempio consideriamo in $\mathbb{R}^4$ i vettori $(2,1,0,0),(0,4,0,1),(0,5,1,0)$ (la prima componente del primo vettore è non nulla mentre la prima componente degli altri due vettori è nulla; il secondo vettore ha la quarta entrata diversa da zero mentre gli altri hanno quarta entrata nulla; infine il terzo vettore ha la terza entrata non nulla e i primi due hanno la terza entrata uguale a 0). Allora una loro combinazione lineare: $\alpha(2,1,0,0) + \beta(0,4,0,1) + \gamma(0,5,1,0) = (2\alpha, \alpha + 4\beta + 5\gamma, \gamma, \beta)$ è uguale a zero se e solo se $\alpha = \beta = \gamma = 0$. Pertanto i vettori $(2,1,0,0),(0,4,0,1),(0,5,1,0)$ sono linearmente indipendenti. Questo esempio ci sarà utile per trovare l'insieme completo di soluzioni di un sistema lineare.

4. Tra tutti gli insiemi di vettori possiamo scegliere anche l'insieme formato da un solo vettore. La domanda è la seguente: quando un vettore v di uno spazio vettoriale V è linearmente indipendente? Per definizione dobbiamo vedere in che modo possiamo scrivere il vettore nullo come sua combinazione lineare: $\lambda v = \mathbf{0}_V$. Ora, abbiamo già visto che se $v \neq \mathbf{0}_V$ allora $\lambda v = \mathbf{0}_V$ se e solo se $\lambda = 0$. Se, invece, $v = \mathbf{0}_V$, allora per ogni $\lambda \in \mathbb{R}$, si ha $\lambda\mathbf{0}_V = \mathbf{0}_V$. Quindi un vettore è linearmente indipendente se e solo se $v \neq \mathbf{0}_V$. Osserviamo inoltre che se un insieme di vettori contiene il vettore nullo, allora è un insieme linearmente dipendente: considerato infatti l'insieme $\{\mathbf{0}_V, v_2, v_3, \dots v_n\}$, la combinazione lineare $(50)\mathbf{0}_V + 0v_2 + 0v_3 + \cdots + 0v_n = \mathbf{0}_V$ ma i suoi coefficienti non sono tutti nulli.

Osservazione 3.1.3 Due vettori sono linearmente dipendenti se e solo se uno dei due è multiplo dell'altro. Infatti se $v_1, v_2 \in V$ spazio vettoriale, allora v_1 e v_2 sono linearmente dipendenti se e solo se esistono $\lambda_1, \lambda_2 \in \mathbb{R}$ non entrambi nulli tali che $\lambda_1 v_1 + \lambda_2 v_2 = \mathbf{0}_V$. Supponiamo che sia $\lambda_1 \neq 0$, allora $v_1 = -(\lambda_2/\lambda_1)v_2$ quindi v_1 è multiplo di v_2.

Osserviamo che sottoinsiemi di insiemi linearmente indipendenti sono ancora linearmente indipendenti come segue facilmente dalla definizione.

La definizione 3.1.1 risponde al quesito introdotto all'inizio della lezione, infatti vale la seguente proposizione.

Proposizione 3.1.4 *In un $\mathbb{R}$-spazio vettoriale V i vettori $u_1, u_2, \dots, u_k$, $k \in \mathbb{N}$, sono linearmente indipendenti se e solo se ogni loro combinazione lineare si scrive in modo unico:*

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = \lambda'_1 u_1 + \lambda'_2 u_2 + \cdots + \lambda'_k u_k, \qquad \lambda_i, \lambda'_i \in \mathbb{R}$$

$$\Updownarrow$$

$$\lambda_1 = \lambda'_1;\ \lambda_2 = \lambda'_2; \dots ; \lambda_k = \lambda'_k.$$

Dimostrazione. "$\Rightarrow$" Per ipotesi sappiamo che (A) i vettori sono linearmente indipendenti e vogliamo mostrare la tesi (B) che la scrittura di una loro combinazione lineare è unica. Vogliamo cioè vedere che $(A) \Rightarrow (B)$. Per farlo mostriamo "l'implicazione opposta": $\neg(B) \Rightarrow \neg(A)$, cioè che non-$(B)$ implica non-(A). Supponiamo che uno stesso vettore di V si scriva in due maniere diverse :

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = \lambda'_1 u_1 + \lambda'_2 u_2 + \cdots + \lambda'_k u_k$$

con qualche $\lambda_i \neq \lambda'_i$, $i = 1, 2, \dots, k$. Allora possiamo scrivere

$$\mathbf{0}_V = v - v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k - (\lambda'_1 u_1 + \lambda'_2 u_2 + \cdots + \lambda'_k u_k)$$

e per la proprietà distributiva del prodotto per scalari rispetto alla somma:

$$\begin{aligned} \mathbf{0}_V &= \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k - \lambda'_1 u_1 - \lambda'_2 u_2 + \cdots - \lambda'_k u_k \\ &= (\lambda_1 - \lambda'_1)u_1 + (\lambda_2 - \lambda'_2)u_2 + \cdots + (\lambda_k - \lambda'_k)u_k. \end{aligned}$$

Ora, almeno uno tra i coefficienti $\lambda_i - \lambda'_i$ è diverso da zero, quindi abbiamo scritto il vettore nullo come una combinazione lineare dei vettori

$\{u_1, u_2, \ldots, u_k\}$ a coefficienti non tutti nulli. Pertanto $\{u_1, u_2, \ldots, u_k\}$ è un insieme di vettori linearmente dipendenti.

"$\Leftarrow$" In questo caso la nostra ipotesi è che ogni vettore che è combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$ lo sia in modo unico. Automaticamente il vettore nullo è loro combinazione lineare (lo è di ogni insieme di vettori): $\mathbf{0}_V = 0u_1 + 0u_2 + \cdots + 0u_k$ e questa scrittura è unica per ipotesi. Quindi abbiamo provato che i vettori $u_1, u_2, \ldots, u_k$ sono linearmente indipendenti. **C. V. D.**

3.2 Basi e dimensione

Siano V uno spazio vettoriale e $I = \{u_1, u_2, \ldots, u_n\}$ un insieme di vettori. Nelle sezioni precedenti abbiamo risolto i seguenti quesiti:

Quesito 1: quando è possibile scrivere ogni vettore di V come combinazione lineare dei vettori di I? Risposta: quando I è insieme di generatori di V.

Quesito 2: quando è possibile scrivere ogni combinazione lineare di vettori di I in modo unico? Risposta: quando I è insieme di vettori linearmente indipendenti.

In questo paragrafo uniremo queste due nozioni: vogliamo che ogni vettore di V si scriva in modo unico come combinazione lineare degli elementi di I. In questa frase il termine "ogni" indica che stiamo cercando un insieme di generatori e il termine "unico" che stiamo cercando un insieme di vettori lineramente indipendenti.

Definizione 3.2.1 *In un $\mathbb{R}$-spazio vettoriale V un insieme (finito) ordinato di vettori $\{b_1, b_2, \ldots, b_s\}$ si dice base (finita) di V se è un insieme di generatori linearmente indipendenti di V.*

Si noti che se V ha una base finita allora è finitamente generato, ma non tutti gli insiemi di generatori sono basi. Mostreremo come, dato un insieme di generatori di uno spazio vettoriale, sia possibile trovarne un sottoinsieme che risulti ancora un insieme di generatori. Cerchiamo di essere più precisi:

Proposizione 3.2.2 *Sia $I = \{u_1, u_2, \ldots, u_k\}$ un insieme di generatori di un $\mathbb{R}$-spazio vettoriale V. Se i vettori $u_1, u_2, \ldots, u_k$ sono linearmente dipendenti esiste un sottoinsieme proprio di I costituito da generatori di V.*

Dimostrazione. Sia

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k = \mathbf{0}_V$$

una relazione di dipendenza tra i vettori $u_1, u_2, \ldots, u_k$, cioè una scrittura del vettore nullo in cui non tutti i coefficienti sono nulli e supponiamo, per fissare le idee, $\lambda_i \neq 0$. Allora sommando il vettore $-\lambda_i u_i$ ad ambedue i membri della precedente relazione si ha:

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_{i-1} u_{i-1} + \lambda_{i+1} u_{i+1} \cdots + \lambda_k u_k = -\lambda_i u_i$$

e, poiché $\lambda_i \neq 0$, possiamo moltiplicare entrambi i membri per $-\frac{1}{\lambda_i}$ ottenendo

$$u_i = -\frac{\lambda_1}{\lambda_i} u_1 - \frac{\lambda_2}{\lambda_i} u_2 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} u_{i+1} - \cdots - \frac{\lambda_k}{\lambda_i} u_k.$$

Sia dunque v un qualunque vettore di V. Allora possiamo scrivere v come combinazione lineare dei vettori $u_1, u_2, \ldots, u_k$, dal momento che questi generano V:

$$v = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_{i-1} u_{i-1} + \beta_i u_i + \beta_{i+1} u_{i+1} \cdots + \beta_k u_k$$

e sostituendo ad u_i l'espressione precedente otteniamo:

$$\begin{aligned} v &= \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_{i-1} u_{i-1} + \beta_{i+1} u_{i+1} \cdots + \beta_k u_k + \\ &+ \beta_i \Big(-\frac{\lambda_1}{\lambda_i} u_1 - \frac{\lambda_2}{\lambda_i} u_2 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} u_{i+1} \cdots - \frac{\lambda_k}{\lambda_i} u_k \Big) \end{aligned}$$

che è una combinazione lineare dei vettori $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots u_k$, infatti:

$$\begin{aligned} v &= \Big(\beta_1 - \beta_i \frac{\lambda_1}{\lambda_i}\Big) u_1 + \Big(\beta_2 - \beta_i \frac{\lambda_2}{\lambda_i}\Big) u_2 + \cdots + \Big(\beta_{i-1} - \beta_i \frac{\lambda_{i-1}}{\lambda_i}\Big) u_{i-1} + \\ &+ \Big(\beta_{i+1} - \beta_i \frac{\lambda_{i+1}}{\lambda_i}\Big) u_{i+1} + \cdots + \Big(\beta_k - \beta_i \frac{\lambda_k}{\lambda_i}\Big) u_k \end{aligned}$$

e quindi $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots u_k$ sono un insieme di generatori di V. **C.V.D.**

Abbiamo dunque dimostrato che, dato un insieme di generatori, se questi sono linearmente dipendenti possiamo eliminarne alcuni ed ottenere ancora un insieme di generatori. Ma se l'insieme di generatori I fosse un insieme di vettori linearmente indipendenti? Quanto detto prima non sarebbe più vero. Infatti se in un $\mathbb{R}$-spazio vettoriale V i vettori $v_1, v_2, \ldots, v_k$ sono linearmente indipendenti esistono dei vettori di v che sono combinazione lineare di $v_1, v_2, \ldots, v_{k-1}, v_k$ ma non di $v_1, v_2, \ldots, v_{k-1}$. Ad esempio v_k non è combinazione lineare di $v_1, v_2, \ldots, v_{k-1}$ (ma bensì di $v_1, v_2, \ldots, v_{k-1}, v_k$, infatti $v_k = 0v_1 + 0v_2 + \cdots + 0v_{k-1} + (1)v_k$). Se fosse: $v_k = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k-1} v_{k-1}$ con $\alpha_i \in \mathbb{R}$, avremmo

$$\mathbf{0}_V = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{k-1} v_{k-1} + (-1)v_k$$

che è una scrittura del vettore nullo come combinazione lineare di $v_1, \ldots, v_k$ in cui almeno un coefficiente è diverso da 0: quello di v_k. Concludendo, un insieme di generatori può essere raffinato se tali vettori sono linearmente dipendenti, ma se si giunge ad un insieme di generatori che è costituito da vettori linearmente indipendenti allora toglierne uno equivale a non riuscire a scrivere più tutti i vettori di V come loro combinazione lineare! Ne deduciamo che la costruzione di una base di uno spazio vettoriale V è legata all'individuazione del minimo numero di generatori di V.

Abbiamo visto quando è possibile eliminare un vettore da un insieme di generatori ottenendo ancora un insieme di generatori, ora vediamo quando è possibile aggiungere un vettore ad un insieme di vettori linearmente indipendenti ottenendo ancora un insieme di vettori linearmente indipendenti.

Osservazione 3.2.3 *In un $\mathbb{R}$-spazio vettoriale V sia dato un insieme di vettori $\{u_1, u_2, \ldots, u_k\}$ linearmente indipendenti. Allora i vettori $u_1, u_2, \ldots, u_k, u$ sono linearmente indipendenti se e solo se u non appartiene al sottospazio $\langle u_1, u_2, \ldots, u_k \rangle$.*

Dimostrazione. "$\Rightarrow$" Supponiamo che $u_1, u_2, \ldots, u_k, u$ siano linearmente indipendenti e che, per assurdo, u appartenga a $\langle u_1, u_2, \ldots, u_k \rangle$, allora $u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k$ con $\alpha_i \in \mathbb{R}$. Sommando $-u$ ai due membri dell'uguaglianza otteniamo $\mathbf{0}_V = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + (-1)u$ e, poiché il coefficiente di u è diverso da zero, i vettori $u_1, u_2, \ldots, u_k, u$ sono linearmente dipendenti, contro l'ipotesi.

"$\Leftarrow$" Supponiamo ora che u non appartenga al sottospazio $\langle u_1, u_2, \ldots, u_k \rangle$ e supponiamo, per assurdo, che $u_1, u_2, \ldots, u_k, u$ siano linearmente dipendenti, pur essendo $u_1, u_2, \ldots, u_k$ linearmente indipendenti. Una relazione di dipendenza per $u_1, u_2, \ldots, u_k, u$ sarà data da

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k + \alpha u = \mathbf{0}_V \tag{3.1}$$

con qualcuno dei numeri reali $\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha$ diverso da zero. Ora α non può essere uguale a 0, poiché allora avremmo una relazione di dipendenza tra $u_1, u_2, \ldots, u_k$ che sono linearmente indipendenti, quindi $\alpha \neq 0$ e $\alpha u = -\alpha_1 u_1 - \alpha_2 u_2 - \cdots - \alpha_k u_k$. Moltiplicando tale uguaglianza per $1/\alpha$, si ottiene u come combinazione lineare di $u_1, u_2, \ldots, u_k$, contro l'ipotesi. **C.V.D.**

Finora quello che abbiamo fatto è stato da una parte, nel caso in cui avessimo un insieme di generatori, mostrare come togliere alcuni vettori e mantenere ancora un insieme di generatori, e dall'altra, nel caso in cui avessimo un insieme di vettori linearmente indipendenti, mostrare come aggiungere altri vettori e mantenere il fatto che fossero linearmente indipendenti. Ma questi due procedimenti hanno un termine? Sì: essi terminano nella individuazione di una base. Riassumiamo quanto detto nel seguente corollario:

Corollario 3.2.4 *Dati uno spazio vettoriale V ed un insieme B di vettori di V, i fatti seguenti sono equivalenti:*

1. *B è una base di V;*

2. *B è un insieme massimale di vettori linearmente indipendenti;*

3. *B è un insieme minimale di generatori di V ;*

(qui massimale e minimale si intendono rispetto all'ordine dato dalle inclusioni come sottoinsiemi di V).

Siamo al risultato culmine della teoria: mostreremo che ogni spazio vettoriale finitamente generato V possiede sempre una base e mostreremo che ogni base di V ha lo stesso numero di elementi, che chiameremo *dimensione* dello spazio vettoriale V.

Teorema 3.2.5 *In un $\mathbb{R}$-spazio vettoriale V sia $I = \{u_1, u_2, \ldots, u_k\}$ un insieme di vettori linearmente indipendenti e sia $G = \{w_1, w_2, \ldots, w_p\}$ un insieme di generatori, con $p, k \in \mathbb{N}$. Allora $k \leq p$, cioè la cardinalità di un insieme di generatori è sempre maggiore della cardinalità di un insieme di vettori linearmente indipendenti o uguale ad essa.*

Dimostrazione. L'idea alla base della dimostrazione è quella di sostituire opportuni elementi di G con elementi di I, ottenendo sempre insiemi di generatori. Alla fine avremo inserito tutti gli elementi di I al posto di (altrettanti) elementi di G, da cui otterremo la conclusione che k (il numero di elementi di I) doveva essere minore o uguale a p (numero di elementi di G). Vediamo i dettagli. Dal momento che i vettori $w_1, w_2, \dots, w_p$ generano V, possiamo esprimere il vettore u_1 come loro combinazione lineare: $u_1 = \alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_p w_p$ e, dal momento che u_1 è non nullo (altrimenti i vettori $u_1, u_2, \dots, u_k$ sarebbero dipendenti), uno tra gli α_i è diverso da zero. Supponiamo, per fissare le idee, $\alpha_1 \neq 0$. Abbiamo allora:

$$w_1 = \frac{1}{\alpha_1} u_1 - \frac{1}{\alpha_1} \alpha_2 w_2 - \dots - \frac{1}{\alpha_1} \alpha_p w_p. \tag{3.2}$$

Ne consegue che $u_1, w_2, \dots, w_p$ sono un insieme di generatori di V. Infatti per ogni vettore $v \in V$ vale $v = \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_p w_p$ con $\beta_i \in \mathbb{R}$ cioè, usando la relazione (3.2), $v = \beta_1(\frac{1}{\alpha_1} u_1 - \frac{1}{\alpha_1}\alpha_2 w_2 - \dots - \frac{1}{\alpha_1}\alpha_p w_p) + \beta_2 w_2 + \dots + \beta_p w_p = \beta_1 \frac{1}{\alpha_1} u_1 + (-\beta_1 \frac{1}{\alpha_1}\alpha_2 + \beta_2) w_2 + \dots + (-\beta_1 \frac{1}{\alpha_1}\alpha_p + \beta_p) w_p$.

Ora scriviamo u_2 come combinazione lineare di $u_1, w_2, \dots, w_p$:

$$u_2 = \gamma_1 u_1 + \gamma_2 w_2 + \dots + \gamma_p w_p \tag{3.3}$$

con $\gamma_i \in \mathbb{R}$. Poiché u_2 non è il vettore nullo, uno tra i γ_i è diverso da 0. Ma non può essere solamente $\gamma_1 \neq 0$: se lo fosse avremmo $u_2 = \gamma_1 u_1$ ma u_1 ed u_2 sono linearmente indipendenti! Dunque almeno uno tra i γ_i con $i \geq 2$ deve essere diverso da zero. Supponiamo, per comodità, che $\gamma_2 \neq 0$. Procediamo allora come prima: sottraiamo $u_2 + \gamma_2 w_2$ ad entrambi i membri della uguaglianza (3.3) e moltiplichiamo per $-\frac{1}{\gamma_2}$:

$$w_2 = \frac{1}{\gamma_2} u_2 - \frac{1}{\gamma_2}\gamma_1 u_1 - \frac{1}{\gamma_2}\gamma_3 w_3 - \dots - \frac{1}{\gamma_2}\gamma_p w_p. \tag{3.4}$$

Procedendo come prima deduciamo che anche $u_1, u_2, w_3, \dots, w_p$ sono dei generatori di V. Ripetendo il ragionamento con u_3 scriviamo $u_3 = \delta_1 u_1 + \delta_2 u_2 + \delta_3 w_3 + \dots + \delta_p w_p$, con $\delta_i \in \mathbb{R}$. Se solo δ_1 e δ_2 fossero diversi da zero avremmo $u_3 = \delta_1 u_1 + \delta_2 u_2$ cioè una relazione di dipendenza lineare tra u_1, u_2 ed u_3, contro l'ipotesi. Si può quindi supporre $\delta_3 \neq 0$. Continuando nello stesso modo, se fosse $k > p$, otterremmo che $u_1, u_2, \dots, u_p$ sono un insieme di generatori, ma allora potremmo esprimere u_{p+1} come loro combinazione lineare:

$u_{p+1} = \epsilon_1 u_1 + \epsilon_2 u_2 + \cdots + \epsilon_p u_p$ con $\epsilon_i \in \mathbb{R}$, cioè avremmo una relazione di dipendenza data da

$$\epsilon_1 u_1 + \epsilon_2 u_2 + \cdots + \epsilon_p u_p + (-1)u_{p+1} + 0u_{p+2} + \cdots + 0u_k = \mathbf{0}_V$$

contro le ipotesi, essendo $u_1, u_2, \ldots, u_k$ linearmente indipendenti. **C.V.D.**

La principale conseguenza del precedente teorema è il seguente corollario:

Corollario 3.2.6 *Ogni base di uno spazio vettoriale ha la medesima cardinalità, cioè lo stesso numero di elementi.*

Dimostrazione. I vettori di una base sono nello stesso tempo linearmente indipendenti e generatori, pertanto se $B_1 = \{c_1, c_2, \ldots, c_l\}$ e $B_2 = \{b_1, b_2, \ldots, b_m\}$ sono due basi di uno spazio vettoriale V, con $l, m \in \mathbb{N}$, allora $l \leq m$ se si pensano $c_1, c_2 \ldots, c_l$ come vettori linearmente indipendenti e $b_1, b_2, \ldots, b_m$ come generatori e $m \leq l$ se si pensano $b_1, b_2, \ldots, b_m$ come linearmente indipendenti e $c_1, c_2 \ldots, c_l$ come generatori. Quindi $l = m$. **C.V.D.**

Questo corollario ci permette di introdurre la seguente definizione:

Definizione 3.2.7 *Si dice* **dimensione di uno spazio vettoriale** *la cardinalità (i.e. il numero di elementi) di una sua qualunque base. Convenzionalmente poniamo la dimensione dello spazio vettoriale banale uguale a 0.*

Teorema 3.2.8 *Ogni spazio vettoriale finitamente generato ammette (almeno) una base. Più precisamente:*

1. *ogni insieme di generatori contiene almeno una base dello spazio;*
2. *ogni insieme di vettori linearmente indipendenti si può completare ad una base.*

Dimostrazione. Per quanto visto, dato un insieme di generatori di uno spazio vettoriale, si può estrarre da questo insieme una base: se i vettori di partenza sono linearmente indipendenti allora sono loro stessi una base; se invece sono linearmente dipendenti per 3.2.2 posso toglierne qualcuno e mantenere il fatto che siano dei generatori. Continuando finché non è più possibile eliminare vettori si ottiene un insieme di generatori linearmente indipendenti, quindi una base.

Analogamente in uno spazio vettoriale V ogni insieme di vettori linearmente indipendenti $u_1, \ldots, u_s$ si può considerare parte di una base: infatti si hanno due possibilità: o $\langle u_1, \ldots, u_s \rangle = V$ e allora $u_1, \ldots, u_s$ sono pure generatori e quindi una base di V, oppure $\langle u_1, \ldots, u_s \rangle$ è contenuto propriamente in V, e dunque esiste $u \in V$, $u \notin \langle u_1, \ldots, u_s \rangle$ e per 3.2.3 $u_1, \ldots, u_s, u$ sono linearmente indipendenti. Così $\langle u_1, \ldots, u_s \rangle \subset \langle u_1, \ldots, u_s, u \rangle$. Continuando in questo modo si ottiene una base di V nel momento in cui si ha un numero di vettori linearmente indipendenti uguale alla dimensione di V. Diremo allora che ogni insieme di vettori linearmente indipendenti di V si può completare in una base di V.

Esempi 3.2.9 1. Consideriamo in $\mathbb{R}^2$ i vettori $e_1 = (1, 0)$ ed $e_2 = (0, 1)$. Essi generano ovviamente $\mathbb{R}^2$, infatti per ogni vettore $v = (a, b)$ di $\mathbb{R}^2$ si ha: $v = (a, b) = a(1, 0) + b(0, 1)$. Inoltre e_1 ed e_2 sono linearmente indipendenti, infatti

$$\alpha e_1 + \beta e_2 = (\alpha, \beta) = (0, 0) \quad \Leftrightarrow \quad \alpha = 0 = \beta.$$

Dunque $B = \{e_1, e_2\}$ è una base di $\mathbb{R}^2$ e $\mathbb{R}^2$ ha pertanto dimensione due. La base B si dice base canonica di $\mathbb{R}^2$. Si noti che le coordinate di un vettore nella base canonica coincidono con le componenti del vettore.

Analogamente, per ogni $n \in \mathbb{N}$ definiamo i seguenti vettori di $\mathbb{R}^n$: $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $\ldots, e_n = (0, 0, \ldots, 0, 1)$. Allora $\{e_1, e_2, \ldots, e_n\}$ è una base di $\mathbb{R}^n$, detta base canonica. La dimensione di $\mathbb{R}^n$ è dunque n. Dato il vettore $(x_1, x_2, \ldots, x_n)$ di $\mathbb{R}^n$ vale: $(x_1, x_2, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n$.

2. Se un sottospazio L di uno spazio vettoriale V di dimensione n ha dimensione n allora coincide con lo spazio ambiente V. Infatti se L ha dimensione n allora esso contiene n vettori linearmente indipendenti: tali vettori, dal momento che le operazioni in L sono quelle di V, risultano linearmente indipendenti anche in V e sono quindi una base di V stesso. Dunque lo spazio da essi generato è $L = V$.

Osservazione 3.2.10 Sia V uno spazio vettoriale finitamente generato. V ammette quindi una base e supponiamo che V abbia dimensione $n \geq 1$. Sia $T \leq V$ un sottospazio non banale di V. Allora anche T è finitamente generato ed ha dimensione $n_1 \leq n$. Mostriamo come costruire una base di T. Sia $v_1 \in T$ un vettore non nullo. Vi sono due possibilità: o $\langle v_1 \rangle = T$ e allora v_1 è un insieme di generatori di T linearmente indipendente, cioè una base, oppure $\langle v_1 \rangle$ è contenuto propriamente in T. Allora esiste in T un vettore v_2 che non è combinazione lineare di v_1 e quindi v_1, v_2 sono linearmente indipendenti in T, ma anche in V, poiché le operazioni in T sono indotte dalle operazioni di V. Allora abbiamo ancora due possibilità: $\langle v_1, v_2 \rangle = T$, e allora $\{v_1, v_2\}$ è un insieme di generatori linearmente indipendenti di T e quindi una sua base, oppure $\langle v_1, v_2 \rangle \subsetneq T$ e possiamo continuare il ragionamento. Il procedimento deve avere una fine poiché i vettori linearmente indipendenti in T sono linearmente indipendenti anche in V, e quindi sono al più n.

Definizione 3.2.11 *Sia V uno spazio vettoriale reale di dimensione n, e sia $\mathbf{u} = \{u_1, \ldots, u_n\}$ una base di V. Allora per ogni vettore $v \in V$ sono univocamente determinati gli scalari $\lambda_i \in \mathbb{R}$ tali che $v = \sum_{i=1}^{n} \lambda_i u_i$; la n-upla $(\lambda_1, \ldots, \lambda_n)$ si dice n-upla delle coordinate di v rispetto alla base* $\mathbf{u}$.

Per indicare che $(\lambda_1, \ldots, \lambda_n)$ è la n-upla di coordinate del vettore v rispetto alla base $\mathbf{u}$, scriveremo $v = (\lambda_1, \ldots, \lambda_n)_{\mathbf{u}}$. Possiamo allora definire l'applicazione

$$f : V \longrightarrow \mathbb{R}^n$$

che associa ad ogni vettore $v \in V$ la n-upla delle sue coordinate nella base $\mathbf{u}$:

$$f(v) = (\lambda_1, \lambda_2, \ldots, \lambda_n).$$

L'applicazione f è ben definita perché $\mathbf{u}$ è una base; è iniettiva, infatti se due vettori hanno le stesse coordinate in una base fissata vuol dire che essi coincidono. Inoltre f è suriettiva, infatti, preso un qualsiasi elemento $(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n$ e posto $w = \sigma_1 u_1 + \sigma_2 u_2 + \cdots + \sigma_n u_n$, w è un elemento di V e $f(w) = (\sigma_1, \ldots, \sigma_n)$. L'applicazione f è dunque biunivoca.

Esempio 3.2.12 Consideriamo l'insieme $L = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - y + z = 0\}$. Si verifichi per esercizio che L è un sottospazio vettoriale di $\mathbb{R}^3$. Vogliamo determinare una base di L. Essendo un sottospazio di uno spazio di dimensione 3, L potrà avere dimensione: 0,1,2,3. Esso non ha dimensione 3 perché allora coinciderebbe con $\mathbb{R}^3$ ma, ad esempio, il vettore $(1, 1, 1) \in \mathbb{R}^3$ non appartiene a L dal momento che le sue coordinate non soddisfano l'equazione di L. Del resto L non è banale poiché contiene almeno il vettore

$(1,2,0) \neq (0,0,0)$. Dunque L avrà dimensione 1 o 2. Abbiamo già notato che $(1,2,0)$ è un vettore di L. Osserviamo che $(0,1,1)$ è un altro vettore che appartiene a L. Dal momento che $(1,2,0)$ e $(0,1,1)$ sono linearmente indipendenti, $\dim L = 2$ e $\{(1,2,0),(0,1,1)\}$ è una base di L.

Osserviamo che, sebbene abbiamo costruito una funzione biiettiva tra ogni spazio vettoriale di dimensione n e $\mathbb{R}^n$, abbiamo definito una base "canonica" solo per $\mathbb{R}^n$.

Esempio 3.2.13 Determiniamo una base dello spazio vettoriale $\mathcal{M}_{2,3}(\mathbb{R})$ delle matrici 2×3 ad entrate reali. Consideriamo allora le seguenti matrici:

$$e_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \, e_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \, e_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$e_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \, e_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \, e_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si verifica facilmente che e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23} sono linearmente indipendenti: ognuna ha una entrata diversa da zero in una posizione in cui tutte le altre matrici hanno una entrata nulla. Inoltre ogni matrice di $\mathcal{M}_{2,3}(\mathbb{R})$ è della forma

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = a_{11}e_{11} + a_{12}e_{12} + a_{13}e_{13} + a_{21}e_{21} + a_{22}e_{22} + a_{23}e_{23}$$

quindi e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23} generano $\mathcal{M}_{2,3}(\mathbb{R})$. La dimensione di $\mathcal{M}_{2,3}(\mathbb{R})$ è pertanto 6. Analogamente si può dimostrare, in generale, che la dimensione di $\mathcal{M}_{n,m}(\mathbb{R})$ è nm e che una base di $\mathcal{M}_{n,m}(\mathbb{R})$ è data dall'insieme delle matrici $\{e_{ij}, i = 1, \dots, n, j = 1, \dots, m\}$ dove e_{ij} è la matrice avente tutte le entrate nulle tranne quella di posto i, j che è uguale a 1.

3.3 Esercizi svolti

Esercizio 3.3.1 Si verifichi che $\{(1,2,3),\ (0,1,0),\ (0,0,1)\}$ è una base di $\mathbb{R}^3$.

Svolgimento. Dobbiamo verificare che

1) i vettori dati sono linearmente indipendenti;

2) i vettori dati generano $\mathbb{R}^3$.

Per provare la lineare indipendenza dobbiamo verificare che, presa una combinazione lineare dei vettori $(1,2,3)$, $(0,1,0)$, $(0,0,1)$, questa è uguale al vettore nullo se e solo se tutti i coefficienti sono nulli. Sia, dunque, $\alpha(1,2,3)+\beta(0,1,0)+\gamma(0,0,1)=(\alpha, 2\alpha+\beta, 3\alpha+\gamma)=(0,0,0)$, cioè

$$\begin{cases} \alpha = 0 \\ 2\alpha+\beta = 0 \\ 3\alpha+\gamma = 0 \end{cases}.$$

Risolvendo il sistema otteniamo: $\alpha=\beta=\gamma=0$. Dunque i vettori $\{(1,2,3),(0,1,0),(0,0,1)\}$ sono linearmente indipendenti. Mostriamo ora che essi generano $\mathbb{R}^3$. Sia (a,b,c) un qualunque vettore di $\mathbb{R}^3$ e cerchiamo $\alpha,\beta,\gamma\in\mathbb{R}$ tali che $(a,b,c)=\alpha(1,2,3)+\beta(0,1,0)+\gamma(0,0,1)=(\alpha, 2\alpha+\beta, 3\alpha+\gamma)$. Otteniamo: $\alpha=a$, $\beta=b-2a$ $\gamma=c-3a$.

Esercizio 3.3.2 In $\mathcal{M}_2(\mathbb{R})$ si considerino le matrici $A=\begin{pmatrix} 3 & 0 \\ 2 & -1 \end{pmatrix}$ e $B=\begin{pmatrix} -5 & 1 \\ 4 & 3 \end{pmatrix}$. Stabilire se A e B sono linearmente indipendenti. Eventualmente completare l'insieme $\{A,B\}$ in una base di $\mathcal{M}_2(\mathbb{R})$.

Svolgimento. Le matrici A e B sono linearmente indipendenti perché non sono una multipla dell'altra. Completare l'insieme $\{A,B\}$ in una base di $\mathcal{M}_2(\mathbb{R})$ significa individuare due elementi $C,D\in\mathcal{M}_2(\mathbb{R})$ tali che $\{A,B,C,D\}$ sia una base di $\mathcal{M}_2(\mathbb{R})$. Procedendo come indicato in 3.2.3 si verifichi che possiamo scegliere $C=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ e $D=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Esercizio 3.3.3 Dato l'insieme $\{(0,1,1),(1,0,-1),(2,3,4),(2,1,0)\}$ di vettori di $\mathbb{R}^3$, stabilire quale delle seguenti affermazioni è vera:

(i) Ogni insieme che contiene quello dato genera $\mathbb{R}^3$.

(ii) Esiste un insieme che contiene quello dato ed è costituito da vettori linearmente indipendenti.

(iii) L'insieme dato è una base di $\mathbb{R}^3$.

Estrarre, se possibile, una base di $\mathbb{R}^3$ dall'insieme dato.

Svolgimento. Ricordiamo innanzitutto che la dimensione di $\mathbb{R}^3$ è 3 quindi 3 è la cardinalità di ogni base di $\mathbb{R}^3$ e quindi il massimo numero di vettori linearmenti indipendenti di $\mathbb{R}^3$. Per questo possiamo immediatamente asserire che le affermazioni (ii) e (iii) sono false. Per convincerci ora della veridicità della affermazione (i) basta verificare che i vettori $(0,1,1)$, $(1,0,-1),(2,3,4),(2,1,0)$ generano $\mathbb{R}^3$. Per questo basta procedere come nell'esercizio 2.4.1. Infine si può mostrare che l'insieme $\{(0,1,1),(1,0,-1),$ $(2,1,0)\}$ è una base di $\mathbb{R}^3$.

Esercizio 3.3.4 Sia $W = \{(2s+t, s-t, s+t, s+2t) \mid s,t \in \mathbb{R}\}$.

(i) Verificare che W è un sottospazio vettoriale di $\mathbb{R}^4$;

(ii) determinare una base $\mathcal{B}$ di W;

(iii) completare $\mathcal{B}$ ad una base $\tilde{\mathcal{B}}$ di $\mathbb{R}^4$.

Svolgimento.

(i) Siano $v = (2s+t, s-t, s+t, s+2t)$ e $w = (2r+p, r-p, r+p, r+2p)$, con $s,t,r,p \in \mathbb{R}$, due elementi di W. Allora $v+w$ è ancora un elemento di W, infatti $v+w = (2s+t+2r+p, s-t+r-p, s+t+r+p, s+2t+r+2p) = (2(s+r)+t+p, (s+r)-(t+p), (s+r)+(t+p), (s+r)+2(t+p))$. Analogamente per ogni $\lambda \in \mathbb{R}$ e per ogni $v \in W$, $\lambda v \in W$. Pertanto W è un sottospazio vettoriale di $\mathbb{R}^4$.

(ii) Individuiamo innanzitutto un insieme di generatori di W. Per questo osserviamo che ogni vettore di W è della forma $(2s+t, s-t, s+t, s+2t) = s(2,1,1,1) + t(1,-1,1,2)$ il che ci consente di affermare che i vettori $(2,1,1,1)$ e $(1,-1,1,2)$ generano W. D'altra parte i vettori $(2,1,1,1)$, $(1,-1,1,2)$ sono linearmente indipendenti (non sono uno multiplo dell'altro!) e quindi individuano una base di W.

(iii) Indicata con $\mathcal{B} = \{(2,1,1,1),(1,-1,1,2)\}$ la base di W individuata in (ii), per ottenere una base $\tilde{\mathcal{B}}$ di $\mathbb{R}^4$ possiamo aggiungere alla base $\mathcal{B}$ i vettori $(0,0,1,0)$ e $(0,0,0,1)$. Si verifica infatti facilmente che i vettori $(2,1,1,1),(1,-1,1,2)$, $(0,0,1,0)$ e $(0,0,0,1)$ sono linearmente indipendenti. Dal momento che $\dim(\mathbb{R}^4) = 4$ questo basta per concludere che essi individuano una base di $\mathbb{R}^4$.

Esercizio 3.3.5 Sia $\mathbb{R}^{\leq 3}[x]$ lo spazio vettoriale reale dei polinomi a coefficienti reali nella variabile x di grado minore o uguale a 3.

(i) Provare che l'insieme dei monomi $\{1, x, x^2, x^3\}$ è una base di $\mathbb{R}^{\leq 3}[x]$. Dedurre che la dimensione di $\mathbb{R}^{\leq 3}[x]$ è 4.

(ii) I vettori $2x^2+1$, $2x+1$, x^3 sono linearmente indipendenti? Completare, se possibile, l'insieme $\{2x^2 + 1, 2x + 1, x^3\}$ in una base di $\mathbb{R}^{\leq 3}[x]$.

(iii) Esistono basi di $\mathbb{R}^{\leq 3}[x]$ costituite da polinomi di grado 3? In caso affermativo esibire un esempio.

(iv) Esistono basi di $\mathbb{R}^{\leq 3}[x]$ costituite da polinomi di grado minore o uguale a 2? In caso affermativo esibire un esempio.

Svolgimento.

(i) Un generico elemento di $\mathbb{R}^{\leq 3}[x]$ è un polinomio $p(x)$ della forma $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 = a_0 \cdot 1 + a_1 \cdot x + a_2 \cdot x^2 + a_3 \cdot x^3$. Dunque $\mathcal{B} = \{1, x, x^2, x^3\}$ è un insieme di generatori di $\mathbb{R}^{\leq 3}[x]$. Del resto una combinazione lineare $a_0 + a_1x + a_2x^2 + a_3x^3$ di $1, x, x^2, x^3$ è uguale al polinomio nullo se e solo se tutti i coefficienti a_0, a_1, a_2, a_3 sono nulli. Questo dimostra che i vettori $1, x, x^2, x^3$ sono linearmente indipendenti e quindi individuano una base di $\mathbb{R}^{\leq 3}[x]$. La dimensione di $\mathbb{R}^{\leq 3}[x]$ è pertanto uguale a 4.

(ii) Sia ora $\alpha(2x^2+1)+\beta(2x+1)+\gamma x^3 = 0$, cioè $\alpha+\beta+2\beta x+2\alpha x^2+\gamma x^3 = 0$. Allora, necessariamente, $\alpha = \beta = \gamma = 0$, pertanto i vettori $2x^2 + 1$, $2x + 1$, x^3 sono linearmente indipendenti. Per completare l'insieme $S = \{2x^2 + 1, 2x + 1, x^3\}$ ad una base di $\mathbb{R}^{\leq 3}[x]$ basta allora aggiungere all'insieme S un polinomio di $\mathbb{R}^{\leq 3}[x]$ che non sia una combinazione lineare di $2x^2 + 1$, $2x + 1$, x^3, ad esempio il polinomio 1 (verificare!). Così l'insieme $\{1, 2x^2 + 1, 2x + 1, x^3\}$ è una base di $\mathbb{R}^{\leq 3}[x]$.

(iii) L'insieme $X = \{x^3, x^3 + x^2, x^3 + x, x^3 + 1\}$ è un esempio di una base di $\mathbb{R}^{\leq 3}[x]$ costituita da polinomi di grado 3. Per rendersi conto che si tratta di una base basta osservare che gli elementi della base $\mathcal{B}$ di (i) si ottengono facilmente come combinazione lineare degli elementi di X. Dunque l'insieme X genera $\mathbb{R}^{\leq 3}[x]$. Dal momento che X ha 4 elementi e che $\dim(\mathbb{R}^{\leq 3}[x]) = 4$, X è una base di $\mathbb{R}^{\leq 3}[x]$.

(iv) Non esiste una base di $\mathbb{R}^{\leq 3}[x]$ costituita da polinomi di grado minore o uguale a 2. Infatti combinando linearmente polinomi di grado minore o uguale a 2 non è possibile ottenere polinomi di grado 3.

Esercizio 3.3.6 Calcolare la dimensione del sottospazio di $\mathbb{R}^4$ generato dai vettori $v_1 = (2,1,0,3)$, $v_2 = (-5,3,4,0)$, $v_3 = (-1,-1,0,2)$, $v_4 = (3,2,0,1)$.

Svolgimento. I vettori v_1 e v_2 sono certo linearmente indipendenti dal momento che non sono uno multiplo dell'altro. Il vettore v_3 è combinazione lineare dei vettori v_1 e v_2? Esistono, cioè, α e $\beta \in \mathbb{R}$ tali che $(-1,-1,0,2) = \alpha(2,1,0,3) + \beta(-5,3,4,0)$? Si tratta di stabilire se esistono α e $\beta \in \mathbb{R}$ tali che

$$\begin{cases} -1 = 2\alpha - 5\beta \\ -1 = \alpha + 3\beta \\ 0 = 4\beta \\ 2 = 3\alpha \end{cases}.$$

Il sistema individuato non ha soluzioni ($\alpha = \frac{2}{3},\ \beta = 0,\ \alpha = -1!$), quindi i vettori v_1, v_2, v_3 sono linearmente indipendenti. Resta, infine, da stabilire se v_4 è combinazione lineare dei vettori v_1, v_2, v_3. Si vede facilmente che $v_4 = v_1 - v_3$. Dunque $\langle v_1, v_2, v_3, v_4 \rangle = \langle v_1, v_2, v_3 \rangle$, pertanto $\dim\langle v_1, v_2, v_3, v_4 \rangle = 3$.

Esercizio 3.3.7 Sia W il sottospazio vettoriale di $\mathbb{R}^5$ definito da $W = \{(x_1, x_2, x_3, x_4, x_5) \mid x_1 - x_2 - 2x_5 = 0,\ x_3 + x_4 + x_5 = 0\}$. Si determini una base $\mathcal{B}$ di W e si calcolino le coordinate del vettore $v = (-4,0,1,1,-2)$ rispetto alla base $\mathcal{B}$.

Svolgimento. Osserviamo innanzitutto che il nostro spazio W non è tutto lo spazio vettoriale ambiente poiché, ad esempio, il vettore $(0,0,0,0,1)$ non soddisfa le due equazioni. Non è neanche lo spazio nullo, poiché il vettore $(1,1,0,0,0)$ appartiene a W. Studiamo le due equazioni che lo caratterizzano: dalla prima otteniamo che $x_1 = x_2 + 2x_5$, dalla seconda $x_3 = -x_4 - x_5$, si vede quindi che scegliendo in modo indipendente x_2, x_4, x_5, possiamo calcolare di conseguenza x_1 e x_3 per ottenere una 5-upla in W. Se scegliamo $x_2 = 1, x_4 = 0, x_5 = 0$ otteniamo il vettore $v_1 = (1,1,0,0,0)$; se scegliamo $x_2 = 0, x_4 = 1, x_5 = 0$ otteniamo il vettore $v_2 = (0,0,-1,1,0)$; se scegliamo $x_2 = 0, x_4 = 0, x_5 = 1$ otteniamo il vettore $v_3 = (2,0,-1,0,1)$. I vettori v_1, v_2, v_3 sono tre vettori di W e sono linearmente indipendenti:

$$\alpha v_1 + \beta v_2 + \gamma v_3 = 0_{\mathbb{R}^5} \quad \Leftrightarrow \quad (\alpha + 2\gamma, \alpha, -\beta - \gamma, \beta, \gamma) = 0_{\mathbb{R}^5}$$

$\Leftrightarrow \quad \alpha = \beta = \gamma = 0$.

Quindi W può avere dimensione 3 o 4. Per determinare una base di W osserviamo che gli elementi di W sono tutti e soli i vettori di $\mathbb{R}^5$ della forma $(x_2 + 2x_5, x_2, -x_4 - x_5, x_4, x_5)$. Ciascuno di questi vettori è esprimibile nel modo seguente:

$$(x_2 + 2x_5, x_2, -x_4 - x_5, x_4, x_5) = x_2(1, 1, 0, 0, 0)+$$

$$+x_5(2, 0, -1, 0, 1) + x_4(0, 0, -1, 1, 0).$$

Pertanto i vettori $v_1 = (1, 1, 0, 0, 0)$, $v_2 = (0, 0, -1, 1, 0)$, $v_3 = (2, 0, -1, 0, 1)$ generano W. Abbiamo così individuato una base di W: $\mathcal{B} = \{v_1, v_2, v_3\}$. In particolare W ha dimensione 3.

Per calcolare le coordinate del vettore $v = (-4, 0, 1, 1, -2)$ rispetto alla base $\mathcal{B}$ dobbiamo determinare α, β, $\gamma \in \mathbb{R}$ tali che $v = \alpha v_1 + \beta v_2 + \gamma v_3$, cioè $(-4, 0, 1, 1, -2) = (\alpha + 2\gamma, \alpha, -\beta - \gamma, \beta, \gamma)$. Otteniamo:

$$\beta = 1, \quad \gamma = -2, \quad \alpha = 0.$$

Pertanto $v = (0, -2, 1)_{\mathcal{B}}$.

3.4 Esercizi proposti

Esercizio 3.4.1 Costruire una base $\mathcal{B}$ di $\mathbb{R}^2$ diversa dalla base canonica e scrivere le coordinate dei vettori della base canonica rispetto alla base $\mathcal{B}$.

Esercizio 3.4.2 Determinare una base $\mathcal{B} = \{v_1, v_2, v_3\}$ di $\mathbb{R}^3$ soddisfacente le seguenti condizioni:

1. le coordinate del vettore $(1, 1, 1)$ rispetto alla base $\mathcal{B}$ sono $(1, 0, 0)$;

2. i vettori v_1, v_2 generano il sottospazio S di $\mathbb{R}^3$: $S = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = 0\}$;

3. le coordinate del vettore $(1, 0, 1)$ rispetto alla base $\mathcal{B}$ sono $(1, 0, 1)$.

La base $\mathcal{B}$ richiesta è unica?

Esercizio 3.4.3 Si considerino i vettori di $\mathbb{R}^3$: $v_1 = (1, 2, 0)$, $v_2 = (1, 1, 1)$, $v_3 = (0, -1, 1)$, $v_4 = (2, 3, 1)$.

1. Stabilire se i vettori v_1, v_2, v_3, v_4 sono linearmente indipendenti.
2. Stabilire se i vettori v_1, v_2, v_3, v_4 generano $\mathbb{R}^3$.
3. Determinare una base del sottospazio di $\mathbb{R}^3$ generato dai vettori v_1, v_2, v_3, v_4.
4. Completare la base trovata in 3. in una base di $\mathbb{R}^3$.

Esercizio 3.4.4 Sia

$$S = \left\{ A = \begin{pmatrix} a & b & c \\ d & f & g \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R}) \mid a+b+d=0,\ d+f+c=0 \right\}.$$

1. Mostrare che S è un sottospazio vettoriale di $\mathcal{M}_{2,3}(\mathbb{R})$.
2. Determinare una base di S.
3. Determinare un sottospazio T di $\mathcal{M}_{2,3}(\mathbb{R})$ tale che $S+T=\mathcal{M}_{2,3}(\mathbb{R})$.

Esercizio 3.4.5 Nell'insieme $V = \mathbb{R}[x,y]$ dei polinomi a coefficienti reali nelle variabili x e y, con le usuali operazioni di somma di polinomi e di prodotto di un polinomio per un numero reale, si consideri il sottoinsieme S dei polinomi di grado minore o uguale a 2.

1. Dopo aver verificato che V è un $\mathbb{R}$-spazio vettoriale e che S è un suo sottospazio, calcolare la dimensione di S ed esibire una sua base $\mathcal{B}$.
2. Calcolare le coordinate del polinomio $x+y-x^2$ nella base $\mathcal{B}$.
3. Mostrare che i polinomi $x-y$, $1+x-y$, $1-xy$ sono linearmente indipendenti.
4. Completare l'insieme $\{x-y, 1+x-y, 1-xy\}$ in una base di V.

Lezione 4

Somma diretta e dimensione di sottospazi

4.1 Somma diretta

Lo scopo di questo paragrafo è il seguente: dati uno spazio vettoriale V ed un suo sottospazio T_1: $T_1 \leq V$, determinare un sottospazio T_2 di V tale che ogni vettore di V si scriva in modo unico come somma di un vettore di T_1 e di un vettore di T_2. Si vuole cioè ottenere una decomposizione di V.

Definizione 4.1.1 *Dati due sottospazi vettoriali T_1 e T_2 di uno spazio vettoriale V, la somma di T_1 e T_2 si dice diretta se $T_1 \cap T_2 = \{\mathbf{0}_V\}$. Detto W lo spazio somma di T_1 e T_2, scriveremo $W = T_1 \oplus T_2$.*

La definizione di somma diretta è la definizione che stavamo cercando, infatti vale la seguente proposizione.

Proposizione 4.1.2 *Sia $W = T_1 \oplus T_2$. Allora ogni vettore $v \in W$ si scrive in modo unico come somma di un elemento di T_1 e di un elemento di T_2: $v = t_1 + t_2$ con $t_1 \in T_1$ e $t_2 \in T_2$.*

Dimostrazione. Essendo $W = T_1 + T_2$, allora se $v \in W$, $v = t_1 + t_2$ con $t_1 \in T_1$ e $t_2 \in T_2$. Tale scrittura è unica: se non lo fosse avremmo infatti $v = t_1 + t_2 = t'_1 + t'_2$, con $t'_1 \in T_1$, $t'_2 \in T_2$ e con $t_1 \neq t'_1$ e/o $t_2 \neq t'_2$ (una delle due disuguaglianze deve essere vera). Ora, sommando $-t'_1 - t_2$ ad entrambi i membri della precedente uguaglianza, si ottiene: $t_1 - t'_1 = t'_2 - t_2$ ma questa è una uguaglianza tra un elemento di T_1 ed un elemento di T_2, quindi $t_1 - t_1' = t_2' - t_2$ è un elemento di $T_1 \cap T_2$. Ma l'intersezione di T_1 e T_2 è il solo vettore nullo, dal momento che la somma di T_1 e T_2 è diretta, quindi $t_1 - t_1' = \mathbf{0}_V = t_2' - t_2$, da cui $t_1 = t_1'$ e $t_2' = t_2$. Dunque l'espressione di v come somma di un elemento di T_1 e di un elemento di T_2 è unica. **CVD**

Se lo spazio vettoriale W possiede due sottospazi T_1 e T_2 tali che $T_1 + T_2 = W$ ma $T_1 \cap T_2 \neq \{\mathbf{0}_V\}$, allora uno stesso vettore di W si può scrivere in due modi diversi come somma di un elemento di T_1 ed uno di T_2 (si veda, a questo proposito, l'esempio 4.1.3). Naturalmente il sottospazio W può anche coincidere con lo spazio ambiente.

Esempio 4.1.3 In $\mathbb{R}^3$ siano dati i due sottospazi $T_1 = \langle (0,1,-1), (1,1,0) \rangle$ e $T_2 = \langle (1,2,-1), (0,0,1) \rangle$. Usando la definizione di spazio vettoriale generato da un insieme di vettori possiamo descrivere esplicitamente gli elementi di ciascun sottospazio: ogni elemento di T_1 è del tipo $\lambda_1(0,1,-1) + \lambda_2(1,1,0)$ al variare di $\lambda_1, \lambda_2 \in \mathbb{R}$, mentre ogni elemento di T_2 è della forma $\eta_1(1,2,-1) + \eta_2(0,0,1)$ al variare di $\eta_1, \eta_2 \in \mathbb{R}$. Per determinare un elemento nella intersezione di T_1 e T_2 si dovranno determinare $\lambda_1, \lambda_2, \eta_1, \eta_2 \in \mathbb{R}$ tali che $\eta_1(1,2,-1) + \eta_2(0,0,1) = \lambda_1(0,1,-1) + \lambda_2(1,1,0)$, vale a dire $(\eta_1, 2\eta_1, -\eta_1 + \eta_2) = (\lambda_2, \lambda_1 + \lambda_2, -\lambda_1)$. Due vettori di $\mathbb{R}^3$ sono uguali se e solo se le loro componenti sono ordinatamente uguali, in questo caso:

$$\begin{cases} \eta_1 = \lambda_2 \\ 2\eta_1 = \lambda_1 + \lambda_2 \\ -\eta_1 + \eta_2 = -\lambda_1. \end{cases}$$

Risolvendo otteniamo: $\eta_1 = \lambda_1 = \lambda_2$ e $\eta_2 = 0$. (Abbiamo trovato tutte le possibili soluzioni del sistema? Perché? Daremo in seguito una risposta precisa a questa domanda). Abbiamo trovato che $T_1 \cap T_2 = \{\eta_1(1,2,-1) \mid \eta_1 \in \mathbb{R}\} = \langle\langle (1,2,-1) \rangle$. In particolare si ha allora che $1(0,1,-1) + 1(1,1,0) = (1,2,-1) \in T_1$ e $1(1,2,-1) + 0(0,0,1) = (1,2,-1) \in T_2$.

Osserviamo che il vettore $(1,3,-1)$ si può scrivere in due modi diversi come somma di un elemento di T_1 e di un elemento di T_2: $(1,3,-1) = (1,3,-2) + (0,0,1) = (0,1,-1) + (1,2,0)$. La somma di T_1 e T_2, infatti, non è diretta.

Osservazione 4.1.4 Supponiamo che V sia uno spazio vettoriale finitamente generato e che T_1 e T_2 siano sottospazi di V tali che $V = T_1 \oplus T_2$. Allora anche T_1 e T_2 sono finitamente generati. Siano dunque $u_1, \ldots, u_k$ e

$v_1, \dots, v_p$ rispettivamente una base di T_1 ed una base di T_2. L'unione delle due basi $u_1, \dots, u_k, v_1, \dots, v_p$ è una base per lo spazio somma V.

Dimostriamolo. Ogni vettore $v \in V$ si scrive come somma di un elemento di T_1 e di un elemento di T_2: $v = t_1 + t_2$, con $t_1 \in T_1$ e $t_2 \in T_2$. D'altra parte $t_1 = \beta_1 u_1 + \dots + \beta_k u_k$ e $t_2 = \gamma_1 v_1 + \dots + \gamma_p v_p$ con γ_i e β_j numeri reali (univocamente determinati). Ne deduciamo che: $v = (\beta_1 u_1 + \dots + \beta_k u_k) + (\gamma_1 v_1 + \dots + \gamma_p v_p) = \beta_1 u_1 + \dots + \beta_k u_k + \gamma_1 v_1 + \dots + \gamma_p v_p$ e quindi $u_1, \dots, u_k, v_1, \dots, v_p$ sono un insieme di generatori di V. Resta da vedere che sono linearmente indipendenti: se vi fosse una relazione di dipendenza tra essi si avrebbe $\beta_1 u_1 + \dots + \beta_k u_k + \gamma_1 v_1 + \dots + \gamma_p v_p = \mathbf{0}_V$ per determinati numeri reali γ_j e β_i non tutti nulli. Se tutti i γ_j fossero eguali a 0 avremmo qualche β_i diverso da zero e quindi una relazione di dipendenza che coinvolge solo $u_1, \dots, u_k$ che sono linearmente indipendenti il che è assurdo. Analogamente otterremmo una contraddizione se tutti i β_i fossero uguali a 0. Dobbiamo quindi concludere che qualche β_i e contemporaneamente qualche γ_j sia diverso da zero, allora, posto

$$w = \beta_1 u_1 + \dots + \beta_k u_k = -\gamma_1 v_1 - \dots - \gamma_p v_p,$$

si avrebbe $w \in T_1 \cap T_2 = \mathbf{0}_V$, quindi $\beta_1 u_1 + \dots + \beta_k u_k = \mathbf{0}_V = -\gamma_1 v_1 - \dots - \gamma_p v_p$. Essendo $u_1, \dots, u_k$ linearmente indipendenti, così come $v_1, \dots, v_p$, si ottiene $\beta_1 = \ldots = \beta_k = \gamma_1 = \ldots = \gamma_p = 0$. È stato assurdo pensare che qualche coefficiente fosse diverso da 0.

Esempio 4.1.5 Dall'osservazione 4.1.4 possiamo dedurre un modo per costruire delle decomposizioni in somma diretta di spazi vettoriali. In un $\mathbb{R}$-spazio vettoriale V di dimensione maggiore di uno prendiamo p vettori linearmente indipendenti $u_1, \dots u_p$, $p \in \mathbb{N}$, $p > 1$, e consideriamo il sottospazio da essi generato: $L = \langle u_1, \dots, u_p \rangle$. Dividiamo l'insieme di vettori $\{u_1, \dots, u_p\}$ in due sottoinsiemi disgiunti che per comodità denoteremo con $\{u_1, \dots, u_h\} \cup \{u_{h+1}, \dots, u_p\} = \{u_1, \dots, u_p\}$, $h \in \mathbb{N}$, $h < p$. Allora si ha $L = \langle u_1, \dots u_h \rangle \oplus \langle u_{h+1}, \dots, u_p \rangle$. Che L sia somma dei due sottospazi indicati deriva dal fatto che ogni suo vettore si scrive come $\lambda_1 u_1 + \dots + \lambda_h u_h + \lambda_{h+1} u_{h+1} + \dots + \lambda_p u_p = (\lambda_1 u_1 + \dots + \lambda_h u_h) + (\lambda_{h+1} u_{h+1} + \dots + \lambda_p u_p)$, con $\lambda_i \in \mathbb{R}$. D'altro canto se ci fosse intersezione non nulla tra i due, avremmo una uguaglianza di vettori data da: $\eta_1 u_1 + \dots + \eta_h u_h = \gamma_{h+1} u_{h+1} + \dots + \gamma_p u_p$ per $\gamma_i, \eta_j \in \mathbb{R}$ non tutti nulli. Ma ecco che dalla uguaglianza precedente si otterrebbe la relazione di dipendenza $\eta_1 u_1 + \dots + \eta_h u_h - \gamma_{h+1} u_{h+1} - \dots - \gamma_p u_p = \mathbf{0}_V$, in contraddizione con l'ipotesi.

Esempio 4.1.6 Consideriamo il seguente problema: in $\mathbb{R}^3$ sia B il sottospazio generato dai vettori $(1,2,3),(0,1,1)$; si determini un sottospazio C tale che $\mathbb{R}^3 = B \oplus C$.
Svolgimento. Ogni sottospazio di uno spazio di dimensione finita ha dimensione finita (Osservazione 3.2.10). In particolare questo è vero per B e C. Si vede subito che B ha dimensione 2 poiché i suoi generatori sono linearmente indipendenti (in quanto diversi da 0 e non multipli) e sono quindi una base di B. Sia allora $u_1, \ldots, u_k$ una base di C. Per quanto visto nell'osservazione 4.1.4 i vettori $(1,2,3),(0,1,1),u_1,\ldots,u_k$ sono linearmente indipendenti. Dal momento che $\mathbb{R}^3$ ha dimensione tre ci sono al più 3 vettori linearmente indipendenti in $\mathbb{R}^3$, quindi $k = 1$. Dunque $C = \langle u_1 \rangle$ e $(1,2,3),(0,1,1),u_1$ devono essere linearmente indipendenti (sappiamo già che i primi due lo sono). Per trovare un siffatto u_1 basterà prendere un vettore di $\mathbb{R}^3$ che non sia combinazione lineare di $(1,2,3),(0,1,1)$, i.e. un vettore non appartenente a $\langle (1,2,3),(0,1,1)\rangle$. Ad esempio $u_1 = (0,0,1)$ e quindi $\langle (1,2,3),(0,1,1)\rangle \oplus \langle (0,0,1)\rangle = \mathbb{R}^3$. Si faccia attenzione che anche $u_1 = (0,1,2)$ va bene, e si ha $\langle (1,2,3),(0,1,1)\rangle \oplus \langle (0,1,2)\rangle = \mathbb{R}^3$, ma $\langle (0,1,2)\rangle \neq \langle (0,0,1)\rangle$

4.2 Dimensione di sottospazi

Presi due sottoinsiemi finiti di un insieme se vogliamo contare il numero di elementi della loro unione possiamo innanzitutto sommare il numero di elementi di uno al numero di elementi dell'altro: in questo modo però gli elementi che stanno nella intersezione vengono contati due volte. Per calcolare allora il numero esatto di elementi dell'unione dobbiamo sottrarre alla somma precedente il numero degli elementi dell'intersezione. Questo è essenzialmente ciò che succede anche per i sottospazi, sostituendo al numero di elementi del sottoinsieme la dimensione del sottospazio e alla unione la somma di sottospazi.

Osservazione 4.2.1 Osserviamo innanzitutto che se in un $\mathbb{R}$-spazio vettoriale V si hanno due sottospazi finitamente generati T_1 e T_2 allora il loro spazio somma è finitamente generato. In effetti se si prendono un insieme di generatori $u_1, \ldots, u_s$ di T_1 ed un insieme di generatori $v_1, \ldots, v_p$ di T_2 allora la loro unione è un insieme di generatori di $T_1 + T_2$ dal momento che ogni vettore di $T_1 + T_2$ si scrive come $v = t_1 + t_2$ con $t_1 \in T_1$ e $t_2 \in T_2$. A loro volta, per ipotesi, $t_1 = \eta_1 u_1 + \cdots + \eta_s u_s$ e $t_2 = \gamma_1 v_1 + \cdots + \gamma_p v_p$, con $\gamma_i, \eta_j \in \mathbb{R}$. Dunque: $v = t_1 + t_2 = (\eta_1 u_1 + \cdots + \eta_s u_s) + (\gamma_1 v_1 + \cdots + \gamma_p v_p) = \eta_1 u_1 + \cdots + \eta_s u_s + \gamma_1 v_1 + \cdots + \gamma_p v_p$.

Teorema 4.2.2 (Formula di Grassmann) *Siano V un $\mathbb{R}$-spazio vettoriale, T_1 e T_2 due suoi sottospazi di dimensione finita, rispettivamente n_1 e n_2. Allora anche l'intersezione $T_1 \cap T_2$ ha dimensione finita e si ha:*

$$\dim(T_1 + T_2) = \dim T_1 + \dim T_2 - \dim(T_1 \cap T_2).$$

Dimostrazione. Certamente $T_1 \cap T_2$ ha dimensione finita in quanto sottospazio vettoriale di uno spazio di dimensione finita (T_1, ad esempio, o T_2). Sia ora $B = \{v_1, \ldots, v_k\}$ una base di $T_1 \cap T_2$. B è un insieme di vettori linearmente indipendenti di T_1 (e di T_2) e quindi si può completare da un lato in una base di T_1: $\mathcal{B}_1 = \{v_1, \ldots, v_k, v'_{k+1}, \ldots, v'_{n_1}\}$, e dall'altro in una base di T_2: $\mathcal{B}_2 = \{v_1, \ldots, v_k, v''_{k+1}, \ldots, v''_{n_2}\}$. Con gli indici sopra introdotti abbiamo $\dim T_1 = n_1$, $\dim T_2 = n_2$ e $\dim T_1 \cap T_2 = k$. Il teorema sarà dimostrato se proveremo che $\mathcal{B}_1 \cup \mathcal{B}_2 = \{v_1, \ldots, v_k, v'_{k+1}, \ldots, v'_{n_1}, v''_{k+1}, \ldots, v''_{n_2}\}$, che ha proprio n_1+n_2-k elementi, è una base di T_1+T_2. Dall'osservazione 4.2.1 sappiamo che $\mathcal{B}_1 \cup \mathcal{B}_2$ è un insieme di generatori di T_1+T_2. Vediamo ora che esso è un insieme di vettori linearmenti indipendenti. Se non lo fosse potremmo scrivere $\mathbf{0}_V = \delta_1 v_1 + \cdots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \cdots + \delta_{n_1} v'_{n_1} + \delta_{n_1+1} v''_{k+1} + \cdots + \delta_{n_1+n_2-k} v''_{n_2}$ per opportuni $\delta_j \in \mathbb{R}$ non tutti nulli. Avremmo cioè $\delta_1 v_1 + \cdots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \cdots + \delta_{n_1} v'_{n_1} = -\delta_{n_1+1} v''_{k+1} - \cdots - \delta_{n_1+n_2-k} v''_{n_2}$. Al primo membro dell'uguaglianza ottenuta abbiamo un vettore di T_1 e al secondo membro un elemento di T_2, quindi un elemento di $T_1 \cap T_2$. Ma se questo vettore appartiene all'intersezione di T_1 e T_2 esso è combinazione lineare dei soli vettori $v_1, \ldots, v_k$ dal momento che $\{v_1, \ldots, v_k\}$ è una base di $T_1 \cap T_2$. Quindi esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tali che $\alpha_1 v_1 + \cdots + \alpha_k v_k = -\delta_{n_1+1} v''_{k+1} - \cdots - \delta_{n_1+n_2-k} v''_{n_2}$ da cui otteniamo $\alpha_1 v_1 + \cdots + \alpha_k v_k + \delta_{n_1+1} v''_{k+1} + \cdots + \delta_{n_1+n_2-k} v''_{n_2} = \mathbf{0}_V$, ma allora $\alpha_1 = \cdots = \alpha_k = \delta_{n_1+1} = \delta_{n_1+n_2-k} = 0$ poiché $\mathcal{B}_2$ è base di T_2 e quindi insieme di vettori linearmente indipendenti.

Resta così : $\delta_1 v_1 + \cdots + \delta_k v_k + \delta_{k+1} v'_{k+1} + \cdots + \delta_{n_1} v'_{n_1} = \mathbf{0}_V$, ma adesso $v_1, \ldots, v_k, v'_{k+1}, \ldots v'_{n_1}$ sono linearmente indipendenti e quindi il solo modo di scrivere il vettore nullo come loro combinazione lineare è attraverso tutti i coefficienti uguali a zero. **C.V.D.**

Esempio 4.2.3 In $\mathbb{R}^3$ consideriamo i due sottospazi $Z_1 = \langle (0,1,0), (1,1,1) \rangle$ e $Z_2 = \langle (1,2,0), (1,1,-1) \rangle$. Z_1 e Z_2 hanno entrambi dimensione 2 e $\dim(Z_1+$

$Z_2) = \dim Z_1 + \dim Z_2 - \dim(Z_1 \cap Z_2)$. Ora, dal momento che $Z_1 + Z_2$ è un sottospazio di $\mathbb{R}^3$, avrà al più dimensione 3, quindi da $\dim(Z_1 + Z_2) = 4 - \dim(Z_1 \cap Z_2)$ si deduce che $Z_1 \cap Z_2$ non può essere lo spazio vettoriale banale. D'altra parte $Z_1 \cap Z_2$ è un sottospazio di Z_1 (e di Z_2) e dunque potrà avere dimensione 1 o 2. Se avesse dimensione 2 allora sarebbe un sottospazio di dimensione due in uno spazio di dimensione 2 cioè coinciderebbe con Z_1 (e con Z_2), si avrebbe quindi: $Z_1 = Z_2 = Z_1 \cap Z_2$, ma questo non è vero, perché $(0, 1, 0) \notin Z_2$, quindi Z_1 ha dimensione 1 e $\dim(Z_1 + Z_2) = 3$, i.e., $Z_1 + Z_2 = \mathbb{R}^3$. Chi è $Z_1 \cap Z_2$? Sappiamo che è uno spazio vettoriale di dimensione 1, quindi per trovarne una base basterà esibire un suo elemento diverso dal vettore nullo. Si tratta quindi di determinare $\eta_1, \eta_2, \epsilon_1, \epsilon_2 \in \mathbb{R}$, per cui $\eta_1(0, 1, 0) + \eta_2(1, 1, 1) = \epsilon_1(1, 2, 0) + \epsilon_2(1, 1, -1)$, cioè: $(\eta_2, \eta_2 + \eta_1, \eta_2) = (\epsilon_2 + \epsilon_1, \epsilon_2 + 2\epsilon_1, -\epsilon_2)$, i.e. $\eta_2 = -\epsilon_2, \epsilon_1 = -2\epsilon_2, \eta_1 = -2\epsilon_2$, ad esempio $\epsilon_2 = 1, \eta_2 = -1, \epsilon_1 = -2, \eta_1 = -2 : (-1, -3, -1) \in Z_1 \cap Z_2$.

4.3 Esercizi svolti

Esercizio 4.3.1 In $\mathcal{M}_2(\mathbb{R})$ (cioè nello spazio vettoriale delle matrici quadrate di ordine 2, a coefficienti reali, che sappiamo avere dimensione 4, si veda l'esempio 3.2.13) si consideri l'insieme $U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{R}) \mid a^2 = d^2 \right\}$. L'insieme U è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$? In caso di risposta negativa, si determini il sottospazio vettoriale generato da U.

Svolgimento. La prima cosa da verificare è che U contenga il vettore nullo. Il vettore nullo in $\mathcal{M}_2(\mathbb{R})$ è la matrice nulla che appartiene a U. Osserviamo che le matrici $\begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix}$ e $\begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$ appartengono ad U. Se U fosse un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$ dovrebbe essere chiuso rispetto alla somma cioè la somma ($+_{\mathcal{M}_2(\mathbb{R})}$) di due vettori di U dovrebbe appartenere anch'essa ad U. Osserviamo allora che se sommiamo le prime due matrici indicate:

$$\begin{pmatrix} -1 & 3 \\ 4 & 1 \end{pmatrix} +_{\mathcal{M}_2(\mathbb{R})} \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4 & 3 \end{pmatrix}$$

la matrice somma non appartiene ad U dal momento che $1^2 \neq 3^2$! Quindi U non è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$. Si noti, però, che se $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in U$

(i.e. $a^2 = d^2$), preso $\alpha \in \mathbb{R}$ allora $\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix} \in U$: infatti se $a^2 = d^2$ allora $(\alpha a)^2 = (\alpha d)^2$.

Il sottospazio generato da U è, per definizione, il più piccolo sottospazio vettoriale che contenga U. Tale sottospazio deve pertanto contenere tutti gli elementi di U ed ogni loro combinazione lineare. Osserviamo che U contiene i vettori $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ e i vettori $C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Lo spazio che stiamo cercando contiene allora le matrici A, B, C, D e tutte le loro combinazioni lineari. Ad esempio conterrà la somma di C e D: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ e la loro differenza $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (-1)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$. Lo spazio generato da U contiene allora le matrici $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ e $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$: ma questi sono quattro vettori linearmente indipendenti di $\mathcal{M}_2(\mathbb{R})$ e quindi una base di $\mathcal{M}_2(\mathbb{R})$ (dal momento che $\mathcal{M}_2(\mathbb{R})$ ha dimensione 4). Pertanto lo spazio generato da U è tutto $\mathcal{M}_2(\mathbb{R})$.

Esercizio 4.3.2 Siano $U = \{(x, y, z) \in \mathbb{R}^3 \mid x+y+z = 0\}$ e $V = \{(x, y, z) \in \mathbb{R}^3 \mid x = y\}$.

(i) Verificare che U e V sono sottospazi vettoriali di $\mathbb{R}^3$.

(ii) Determinare una base di U ed una base di V.

(iii) Determinare $U \cap V$ e $U + V$.

(iv) Completare la base di $U \cap V$ ad una base di $\mathbb{R}^3$.

Svolgimento. A questo punto dello studio, lo studente dovrebbe essere in grado di risolvere la parte (i) senza colpo ferire, se non è questo il caso ....si riparta da 2.1.2.

(ii) Gli elementi di U sono i vettori di $\mathbb{R}^3$ della forma $(x, y, -x-y) = x(1, 0, -1)+y(0, 1, -1)$ con $x, y \in \mathbb{R}$. Pertanto l'insieme $\{(1, 0, -1), (0, 1, -1)\}$ genera U. D'altra parte i vettori $(1, 0, -1)$, $(0, 1, -1)$ sono linearmente indipendenti e quindi individuano una base di U.

Analogamente gli elementi di V sono tutti e soli della forma (a, a, z) con a e z in $\mathbb{R}$, cioè $a(1, 1, 0) + z(0, 0, 1)$. Così $\{(1, 1, 0), (0, 0, 1)\}$ genera V ed è una sua base dal momento che i vettori $(1, 1, 0)$ e $(0, 0, 1)$ sono linearmente indipendenti.

(iii) Il sottospazio $U + V$ contiene tutti gli elementi di U e tutti gli elementi di V. In particolare contiene i vettori $(1, 0, -1), (0, 1, -1)$, $(0, 0, 1)$. Questi tre vettori sono linearmente indipendenti e generano quindi $\mathbb{R}^3$. Di conseguenza $U+V$ coincide con $\mathbb{R}^3$ e l'insieme $\{(1, 0, -1), (0, 1, -1), (0, 0, 1)\}$ è una sua base (come pure $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ o qualsiasi altra base di $\mathbb{R}^3$). Il teorema 4.2.2 ci consente ora di calcolare la dimensione di $U \cap V$: $\dim(U \cap V) = \dim U + \dim V - \dim(U + V) = 2 + 2 - 3 = 1$. Pertanto per determinare una base di $U \cap V$ sarà sufficiente determinare un vettore non nullo appartenente sia ad U che a V. Ad esempio, il vettore $(1, 1, -2)$ soddisfa sia l'equazione di U che quella di V e appartiene pertanto alla loro intersezione. Dunque $U \cap V = \langle (1, 1, -2) \rangle$.

(iv) Per completare $\{(1, 1, -2)\}$ in una base di $\mathbb{R}^3$ possiamo prendere, ad esempio, i vettori $(0, 1, 0)$ e $(0, 0, 1)$. Perché? Basta verificare che l'insieme $\{(1, 1, -2), (0, 1, 0), (0, 0, 1)\}$ è un insieme di vettori linearmente indipendenti.

Esercizio 4.3.3 In $\mathbb{R}^{\leq 3}[x]$ si considerino i polinomi $p_1 = x^2 - 1$, $p_2 = 3$, $p_3 = 2x^2 + 1$, $q_1 = x^3$, $q_2 = 5$, $q_3 = x + 1$. Siano S e T i sottospazi vettoriali di $\mathbb{R}^{\leq 3}[x]$ generati rispettivamente da p_1, p_2, p_3 e q_1, q_2, q_3. Determinare:

(i) la dimensione di S e la dimensione di T;

(ii) la dimensione di $S \cap T$ e $S + T$ ed una base di ciascuno di essi.

Svolgimento.

(i) S è il sottospazio vettoriale di $\mathbb{R}^{\leq 3}[x]$ generato dai polinomi $p_1 = x^2 - 1$, $p_2 = 3$, $p_3 = 2x^2 + 1$. Cerchiamo di stabilire se questi sono o meno linearmente indipendenti. I polinomi p_1 e p_2 sono certamente linearmente indipendenti dal momento che p_1 ha grado 2 e p_2 ha grado 0. Il polinomio p_3 è combinazione lineare di p_1 e p_2? Esistono, in altre parole, α e β in $\mathbb{R}$ tali che sia $p_3 = 2x^2 + 1 = \alpha(x^2 - 1) + \beta(3) = 3\beta - \alpha + \alpha x^2$? In effetti basta prendere $\alpha = 2$ e $\beta = 1$. Dunque p_3 è combinazione lineare di p_1 e p_2 e $S = \langle p_1, p_2 \rangle$ ha dimensione 2.

Procedendo in modo analogo si verifica che i polinomi q_1, q_2 e q_3 sono linearmente indipendenti e che, quindi, T ha dimensione 3.

(ii) A questo punto sappiamo già che la somma di S e T non può essere diretta, cioè che la loro intersezione non può essere banale. Infatti

$$\dim S \cap T = \dim S + \dim T - \dim(S+T) = 2 + 3 - \dim(S+T) \geq$$
$$\geq 2 + 3 - 4 = 1$$

dal momento che $S+T$ è un sottospazio di $\mathbb{R}^{\leq 3}[x]$ e ha pertanto dimensione minore o uguale a quella di $\mathbb{R}^{\leq 3}[x]$ cioè 4. Determiniamo $S \cap T$: un elemento di $S \cap T$ è un polinomio che si scrive contemporaneamente come combinazione lineare di p_1 e p_2 e di q_1, q_2, q_3:

$$a(x^2 - 1) + b(3) = \alpha x^3 + \beta(5) + \gamma(x + 1).$$

Otteniamo dunque: $\alpha = a = \gamma = 0$ e $5\beta = 3b$. Di conseguenza, gli unici polinomi che appartengono sia a T che a S sono i polinomi di grado 0: $S \cap T = \langle 1 \rangle$. Otteniamo allora immediatamente che $\dim(S + T) = 4$ cioè $S + T = \mathbb{R}^{\leq 3}[x]$.

Esercizio 4.3.4 In $V = \mathbb{R}^2$ consideriamo i sottospazi vettoriali $T_1 = \langle (1, 1) \rangle$ e $T_2 = \langle (1, 2) \rangle$. Si vede immediatamente che T_1 e T_2 non hanno vettori in comune diversi dal vettore nullo. Inoltre l'insieme $\{(1, 1), (1, 2)\}$ genera $\mathbb{R}^2$. Dunque: $\mathbb{R}^2 = T_1 \oplus T_2$. Sia ora $S = \langle (1, 3) \rangle$. Allora, come prima, $\mathbb{R}^2 = T_1 \oplus S$, ma $S \neq T_2$. Inoltre vale anche: $\mathbb{R}^2 = S \oplus T_2$ e questo mostra che non solo S e T_2 non hanno vettori in comune, ma addirittura sono talmente diversi da generare tutto lo spazio in somma diretta. D'altro canto non avendo S e T_2 vettori in comune ed essendo lo spazio ambiente $\mathbb{R}^2$, così deve essere. Da questo esercizio risulta chiaro che dato $T \leq V$ (con V di dimensione finita) esiste, ma non è unico, un sottospazio $S \leq V$ tale che $T \oplus S = V$.

Esercizio 4.3.5 Consideriamo i seguenti sottospazi vettoriali di $\mathbb{R}^4$: $S = \{(x, y, z, t) \in \mathbb{R}^4 \mid 3x + y + 2z = 0\}$ e $T = \langle (2, 1, 0, -5), (3, 4, 0, 0) \rangle$. Stabilire se la somma di S e T è diretta e calcolare la dimensione di $S + T$.

Svolgimento. La somma di due sottospazi S e T di uno spazio vettoriale V è diretta se $S \cap T = \{\mathbf{0}_V\}$. Determiniamo dunque l'intersezione di S e T. Un elemento generico di T è un elemento della forma $v = (2\alpha + 3\beta, \alpha + 4\beta, 0, -5\alpha)$ con $\alpha, \beta \in \mathbb{R}$. Il vettore v appartiene a S se e solo se le sue coordinate soddisfano l'equazione di S: $6\alpha + 9\beta + \alpha + 4\beta = 0$, cioè $7\alpha + 13\beta = 0$. Dunque il vettore $v = 13(2, 1, 0, -5) - 7(3, 4, 0, 0) = (5, -15, 0, -65)$ appartiene a $S \cap T$, così come ogni suo multiplo. Pertanto la somma di S e T non è diretta e $S \cap T = \langle v \rangle$. Si ha dunque $\dim(S + T) = \dim S + \dim T - \dim S \cap T = 3 + 2 - 1 = 4$, cioè $S + T = \mathbb{R}^4$.

4.4 Esercizi proposti

Esercizio 4.4.1 Costruire due sottospazi S e T di $\mathbb{R}^3$ la cui somma non sia diretta e sia uguale ad $\mathbb{R}^3$.

Esercizio 4.4.2 Sia $S = \langle (1,1,1,1), (1,0,-1,2), (-1,0,-1,1), (1,2,1,3) \rangle$.

1. Determinare la dimensione di S.
2. Determinare, se possibile, un sottospazio T di $\mathbb{R}^4$ tale che $S \oplus T = \mathbb{R}^4$.
3. Determinare, se possibile, un sottospazio non banale W di $\mathbb{R}^4$ tale che la somma di S e W sia diretta e propriamente contenuta in $\mathbb{R}^4$.
4. Determinare, se possibile, un sottospazio V di $\mathbb{R}^4$ tale che la somma di S e V non sia diretta e sia uguale a $\mathbb{R}^4$.

Esercizio 4.4.3 Si considerino i seguenti sottospazi vettoriali di $\mathcal{M}_2(\mathbb{R})$:

$$U = \left\{ \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}, \quad V = \left\{ \begin{pmatrix} d & 0 \\ d & d \end{pmatrix} \mid d \in \mathbb{R} \right\}.$$

1. Determinare una base di U ed una base di V.
2. Determinare una base di $U + V$ ed una base di $U \cap V$.
3. Completare la base di $U + V$ trovata in 2. in una base $\mathcal{B}$ di $\mathcal{M}_2(\mathbb{R})$.
4. Determinare le coordinate del vettore $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ nella base $\mathcal{B}$.

Esercizio 4.4.4 Sia V un $\mathbb{R}$ spazio vettoriale di dimensione 3 e sia $\{u, v, w\}$ una base di V. Siano $S = \langle u - v, v - w \rangle$ e $T = \langle u + w, 2u - w \rangle$.

1. Calcolare la dimensione di S e la dimensione di T.
2. Determinare $S + T$ e $S \cap T$ ed esibire una base per ciascuno di essi.
3. Completare la base trovata di $S \cap T$ in una base di V.

Esercizio 4.4.5 In $\mathbb{R}^{\leq 2}[x]$ si considerino i sottospazi $S = \langle 1 + x, 1 - x^2 \rangle$ e $T = \langle 2 + x + x^2, x + x^2 \rangle$.

1. Determinare $S + T$ e $S \cap T$.
2. Stabilire se ogni vettore di $\mathbb{R}^{\leq 2}[x]$ si scrive come somma di un vettore di S e di un vettore di T e, in caso affermativo, se tale scrittura è unica.

Lezione 5

Applicazioni lineari e matrici

5.1 Applicazioni lineari

Cos'è una applicazione tra due spazi vettoriali? Richiedere che sia solo un'applicazione tra i due insiemi di vettori sarebbe riduttivo: le strutture in gioco sono più ricche e le applicazioni debbono rispettare queste strutture.

Definizione 5.1.1 *Una applicazione* $L : V \to W$ *tra due* $\mathbb{R}$*-spazi vettoriali* V, W *si dirà lineare se*
- $\forall\ v_1, v_2 \in V$, $L(v_1 +_V v_2) = L(v_1) +_W L(v_2)$.
- $\forall\ v \in V$, $\forall \alpha \in \mathbb{R}$, $L(\alpha v) = \alpha L(v)$.

Questa è la definizione di applicazione lineare. Si noti che dalla definizione segue subito che, per qualunque vettore v di V, $L(\mathbf{0}_V) = L(0v) = 0L(v) = \mathbf{0}_W$. Si noti anche che l'opposto di un vettore v viene mandato nell'opposto del vettore $L(v)$: $L((-1)v) = (-1)L(v)$ che è l'opposto in W del vettore $L(v)$.

Ricordiamo che, data una applicazione qualsiasi f tra due insiemi X, Y, $f : X \to Y$ (quindi X è il dominio dell'applicazione f e Y il suo codominio), per ogni sottoinsieme X_1 di X possiamo definire un sottoinsieme di Y ponendo:

$$f(X_1) = \{y \in Y \mid \exists\, x \in X_1 \text{ tale che } f(x) = y\}.$$

Tale insieme si dirà *insieme immagine di* X_1 *tramite* f . Analogamente se Y_1 è un sottoinsieme di Y possiamo definire il seguente sottoinsieme di X:

$$f^{-1}(Y_1) = \{x \in X \mid f(x) \in Y_1\}$$

(attenzione: si tratta solo di una notazione, NON stiamo richiedendo che la funzione f sia invertibile); $f^{-1}(Y_1)$ si dirà *immagine inversa (o controimmagine o antiimmagine) di* Y_1 *tramite* f.

Proposizione 5.1.2 *Data una applicazione lineare* $L : V \to W$ *tra due* $\mathbb{R}$-*spazi vettoriali* V, W *se* $V_1 \leq V$ *è un sottospazio vettoriale di* V *allora la sua immagine tramite* L, $L(V_1)$, *è un sottospazio vettoriale di* W. *Se* $W_1 \leq W$ *è un sottospazio vettoriale di* W *allora la sua controimmagine tramite* L, $L^{-1}(W_1)$, *è un sottospazio di* V.

Dimostrazione. Dimostriamo la prima affermazione: $L(V_1) = \{w \in W \mid \exists\, v \in V_1 \text{ tale che } L(v) = w\}$ è un sottospazio di W. Per fare questo dobbiamo usare la caratterizzazione di sottospazio che abbiamo dato in 2.1.2. Cioè, presi w_1 e w_2 in $L(V_1)$ (che è in ogni caso un insieme di vettori di W), dobbiamo vedere innanzitutto che $w_1 +_W w_2 \in L(V_1)$. Per definizione $w = w_1 +_W w_2 \in L(V_1)$ se esiste un vettore $v \in V_1$ che ha w come immagine mediante L, cioè: $L(v) = w$. Ora: $w_1, w_2 \in L(V_1)$ quindi esistono $v_1, v_2 \in V_1$ tali che $L(v_1) = w_1$ e $L(v_2) = w_2$, pertanto $w_1 +_W w_2 = L(v_1) +_W L(v_2)$; per la linearità di L si ha che $L(v_1) +_W L(v_2) = L(v_1 +_V v_2)$, ma V_1 è un sottospazio di V e quindi $v_1 +_V v_2 \in V_1$ e allora $w_1 +_W w_2 = L(v_1 +_V v_2)$ sta in $L(V_1)$ poiché è l'immagine di $v_1 +_V v_2 \in V_1$. Analogamente, se $\alpha \in \mathbb{R}$ e $w \in L(V_1)$, allora esiste $v \in V_1$ tale che $L(v) = w$, quindi $\alpha w = \alpha L(v)$ ed essendo L lineare $\alpha w = L(\alpha v)$, ma V_1 è un sottospazio di V, pertanto $\alpha v \in V_1$. Dunque αw è immagine del vettore $\alpha v \in V_1$. Quindi $L(V_1)$ è un sottospazio vettoriale di W.

Dimostriamo la seconda affermazione: siano allora $v_1, v_2 \in L^{-1}(W_1)$, vale a dire $L(v_1) \in W_1$ e $L(v_2) \in W_1$; $v_1 +_V v_2$ appartiene ancora a $L^{-1}(W_1)$? Cioè $L(v_1 +_V v_2) \in W_1$? In effetti L è lineare, quindi $L(v_1 +_V v_2) = L(v_1) +_W L(v_2)$, ma sia $L(v_1)$ che $L(v_2)$ stanno in W_1 ed essendo W_1 un sottospazio, anche la loro somma appartiene a W_1. Per il prodotto per scalari: se $\alpha \in \mathbb{R}$ e $v \in L^{-1}(W_1)$ allora $\alpha v \in L^{-1}(W_1)$ se $L(\alpha v) \in W_1$, ma, ancora, L è lineare e dunque: $L(\alpha v) = \alpha L(v)$. Dal momento che $L(v) \in W_1$ per ipotesi e W_1 è un sottospazio, $\alpha L(v) \in W_1$. **C.V.D.**

Alcuni dei sottospazi costruiti nella proposizione precedente ci interesseranno più degli altri. In particolare l'insieme immagine mediante L del dominio V di L si chiama **immagine di L** e si indica con $\mathrm{Im}L$: $L(V) = \mathrm{Im}L$. Osserviamo che l'applicazione L è suriettiva se e solo se $\mathrm{Im}L = W$.

Consideriamo la controimmagine del vettore nullo di W: $L^{-1}(\{\mathbf{0}_W\})$. Questo è il sottospazio degli elementi di V che sono mandati mediante L nel vettore nullo di W. Tale sottospazio di V si chiama **nucleo** dell'applicazione lineare L e si indica con $\mathrm{Ker}L$:

$$\mathrm{Ker}L = \{v \in V \mid L(v) = \mathbf{0}_W\}.$$

Osserviamo che $\mathrm{Ker}L$ contiene sempre $\mathbf{0}_V$.

La Proposizione 5.1.2 afferma, in particolare, che il nucleo e l'immagine di una applicazione lineare L sono sottospazi vettoriali rispettivamente del dominio e del codominio di L.

Abbiamo visto il legame tra l'immagine di una applicazione lineare e la sua suriettività. Che cosa possiamo dire sulla iniettività? Sappiamo che una funzione $f : X \to Y$ è iniettiva se per ogni coppia di elementi distinti di X $x_1 \neq x_2$ si ha $f(x_1) \neq f(x_2)$. La medesima definizione vale ovviamente per applicazioni lineari iniettive. Come possiamo caratterizzare una applicazione lineare iniettiva? Vale il seguente risultato:

Proposizione 5.1.3 *Data una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W, L è iniettiva se e solo se* $\mathrm{Ker}L$ *è il sottospazio banale di V.*

Dimostrazione. "$\Rightarrow$" Sia L iniettiva. Sappiamo che il sottospazio $\mathrm{Ker}L$ è costituito dai vettori di V che sono mandati mediante L nel vettore nullo di W. Per definizione di applicazione lineare tale spazio contiene sempre il vettore nullo di V. Se contenesse un altro vettore $v \in V$, $v \neq \mathbf{0}_V$, vorrebbe dire $L(v) = L(\mathbf{0}_V) = \mathbf{0}_W$ e questo non è possibile perché la funzione è iniettiva.

"$\Leftarrow$" Sia $\mathrm{Ker}L = \{\mathbf{0}_V\}$ e supponiamo che la funzione non sia iniettiva cioè che esistano due vettori $v_1, v_2 \in V$, $v_1 \neq v_2$, tali che $L(v_1) = L(v_2)$ cioè $L(v_1) - L(v_2) = \mathbf{0}_W$, ma L è lineare dunque $L(v_1 - v_2) = L(v_1) - L(v_2) = \mathbf{0}_W$, pertanto $v_1 - v_2 \in \mathrm{Ker}L$. Essendo $v_1 \neq v_2$, $v_1 - v_2 \neq \mathbf{0}_V$ e questo non è possibile perché $\mathrm{Ker}L = \{\mathbf{0}_V\}$. La funzione L è quindi iniettiva. **C.V.D.**

Riflessione. Una definizione equivalente di applicazione iniettiva $f : X \to Y$ consiste nel dire che l'antiimmagine di ogni elemento di $\mathrm{Im}f$ è costituita da uno ed un solo elemento del dominio. Data un'applicazione lineare $L : V \to W$, il vettore $\mathbf{0}_W$ appartiene sempre all'immagine di L e la proposizione precedente mostra che L è iniettiva se e solo se l'antiimmagine del vettore nullo $\mathbf{0}_W$ è costituita solamente da $\mathbf{0}_V$. Per verificare l'iniettività di un'applicazione lineare L dunque non occorre considerare l'antiimagine di ogni vettore dell'immagine ma solo di $\mathbf{0}_W$.

Quest'ultima proposizione ci aiuta anche a risolvere il seguente esercizio:

Esercizio 5.1.4 *Data un'applicazione lineare $L : V \to W$ e dato un vettore $w \in W$, in che modo possiamo descrivere $L^{-1}(\{w\})$, cioè l'insieme di tutti i vettori di V la cui immagine mediante L è w?*

Soluzione Se $w = \mathbf{0}_W$ allora $L^{-1}(\{w\}) = \mathrm{Ker}L$.

Se $w \neq \mathbf{0}_W$? Se $w \notin \mathrm{Im}L$ la risposta è facile: $L^{-1}(\{w\}) = \emptyset$.

Se, invece, $w \in \mathrm{Im}L$ esiste sicuramente un elemento $v \in V$ tale che $L(v) = w$. Si ha allora:

$$L^{-1}(\{w\}) = v + \mathrm{Ker}L. \tag{5.1}$$

La scrittura $v + \mathrm{Ker}L$ indica l'insieme dei vettori della forma $v + k$ con $k \in \mathrm{Ker}L$. Per dimostrare l'uguaglianza tra i due insiemi in (5.1) dobbiamo mostrare che vale la doppia inclusione $L^{-1}(\{w\}) \subseteq v + \mathrm{Ker}L$ e $v + \mathrm{Ker}L \subseteq L^{-1}(\{w\})$. Dato $v+k \in v+\mathrm{Ker}L$ la sua immagine è $L(v+k)$ che, per linearità, coincide con $L(v) + L(k) = w + \mathbf{0}_W = w$, dunque $v + k \in L^{-1}(\{w\})$; così abbiamo visto che $v+\mathrm{Ker}L \subseteq L^{-1}(\{w\})$. Per l'altra inclusione prendiamo un vettore $s \in L^{-1}(\{w\})$ e consideriamo la somma di s con l'opposto di v: $s-v$. Applichiamo L: $L(s - v) = L(s) - L(v) = w - w = \mathbf{0}_W$, cioè $s - v \in \mathrm{Ker}L$. Così, posto $k = s - v$, si ha $s = v + k \in v + \mathrm{Ker}L$; vale dunque l'inclusione $L^{-1}(\{w\}) \subseteq v + \mathrm{Ker}L$.

Osservazione 5.1.5

i) Nella costruzione di $L^{-1}(\{w\}) = v + \mathrm{Ker}L$ abbiamo scelto un vettore v tale che $L(v) = w$. Se ne avessimo scelto un altro? Il risultato sarebbe stato lo stesso. Infatti scelto $\overline{v} \neq v$ tale che $L(\overline{v}) = w$, avremmo trovato $L^{-1}(\{w\}) = \overline{v} + \mathrm{Ker}L$, quindi $\overline{v} + \mathrm{Ker}L = v + \mathrm{Ker}L$.

ii) Che struttura ha $v + \mathrm{Ker}L$? Non possiamo aspettarci molto. Se $v \notin \mathrm{Ker}L$ allora nessuna delle somme $v + k$, $k \in \mathrm{Ker}L$, può essere il vettore nullo, quindi $v + \mathrm{Ker}L$ non è uno spazio vettoriale (se fosse $v + k = \mathbf{0}_V$ si avrebbe $v = -k \in \mathrm{Ker}L$, assurdo). L'altra possibilità è che $v \in \mathrm{Ker}L$, ma allora $v + \mathrm{Ker}L = \mathrm{Ker}L$ è un sottospazio. (Per mostrare l'uguaglianza basta osservare che ogni elemento z di Ker$L$ si può scrivere come $z = v + (z - v)$ ove $-v \in$ Ker$L$ per ipotesi e dunque $z - v \in$ Ker$L$).

iii) Data una applicazione lineare $L : V \to W$ l'antiimmagine tramite L di un vettore $w \in$ ImL è data da un vettore particolare $v \in V$ che soddisfi la condizione $L(v) = w$ e da tutti i vettori che si ottengono sommando v agli elementi di Ker$L$. Cosicché se L è iniettiva allora $L^{-1}(\{w\})$ è costituito da un solo vettore!

iv) Sia V uno spazio vettoriale di dimensione finita n e sia $\{v_1, v_2, \dots, v_n\}$ una sua base. Consideriamo un altro spazio vettoriale W e n vettori qualsiasi $w_1, w_2, \dots, w_n$ di W (i vettori $w_1, w_2, \dots, w_n$ possono essere scelti del tutto arbitrariamente: possono essere tutti uguali, tutti uguali al vettore nullo, tutti diversi, etc.). Possiamo allora costruire una applicazione lineare $f : V \longrightarrow W$ semplicemente ponendo $f(v_1) = w_1, f(v_2) = w_2, \dots, f(v_n) = w_n$. Non solo: esiste un'unica applicazione lineare $f : V \longrightarrow W$ tale che $f(v_i) = w_i$ con $i = 1, \dots, n$.

Come si costruisce f? Ogni vettore v di V si scrive in modo unico come combinazione lineare dei vettori $v_1, v_2, \dots, v_n$: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Se vogliamo che f sia lineare dovrà dunque essere

$$f(v) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n. \tag{5.2}$$

La condizione (5.2) definisce una funzione lineare. Siano infatti $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e $t = \gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_n v_n$, con $\lambda_i, \gamma_i \in \mathbb{R}$, due elementi di V. Allora la loro somma si scrive in modo unico come $t + v = (\gamma_1 + \lambda_1)v_1 + (\gamma_2 + \lambda_2)v_2 + \cdots + (\gamma_n + \lambda_n)v_n$. E dunque $f(t + v) = (\gamma_1 + \lambda_1)w_1 + (\gamma_2 + \lambda_2)w_2 + \cdots + (\gamma_n + \lambda_n)w_n = (\gamma_1 w_1 + \gamma_2 w_2 + \cdots + \gamma_n w_n) + (\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n) = f(t) + f(v)$. Inoltre, per ogni $\alpha \in \mathbb{R}$, $\alpha v = \alpha\lambda_1 v_1 + \alpha\lambda_2 v_2 + \cdots + \alpha\lambda_n v_n$ e $f(\alpha v) = \alpha\lambda_1 w_1 + \alpha\lambda_2 w_2 + \cdots + \alpha\lambda_n w_n = \alpha(\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n) = \alpha f(v)$. Quindi la applicazione f è lineare. L'unicità di f dipende dal fatto che la condizione $f(v_i) = w_i$ fissa le immagini degli elementi di una base e, per linearità, determina univocamente l'immagine di qualsiasi elemento di V. Osserviamo che il dato delle immagini di un insieme di vettori linearmente indipendenti di V che non sia una base non caratterizza univocamente un'applicazione lineare.

5.2 Struttura dimensionale

Per il momento abbiamo solo informazioni qualitative riguardo ad una applicazione lineare. Adesso vogliamo dare qualche informazione 'quantitativa'. Supponiamo quindi di avere una applicazione lineare $L : V \to W$ e supponiamo inoltre che V abbia dimensione finita e che $v_1, \ldots, v_n$ sia una sua base. Allora le immagini dei vettori della base, $L(v_1), \ldots, L(v_n)$, sono un insieme di generatori di $\mathrm{Im}L$. Infatti, sia $w \in \mathrm{Im}L$; questo vuol dire che esiste un vettore $v \in V$ tale che $L(v) = w$, allora $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, essendo $v_1, \ldots, v_n$ una base di V. Applicando L, per linearità, si ottiene: $w = L(v) = \lambda_1 L(v_1) + \cdots + \lambda_n L(v_n)$ e quindi, $L(v_1), \ldots, L(v_n)$ sono un insieme di generatori di $\mathrm{Im}L$. Attenzione: NON stiamo dicendo che $L(v_1), \ldots, L(v_n)$ sono una base dell'immagine cioè che sono linearmente indipendenti! Ma dal loro insieme (come da ogni insieme di generatori) si potrà estrarre una base dell'immagine, il che consentirà di calcolare la dimensione di $\mathrm{Im}L$. Il risultato seguente risponderà a tutti i nostri quesiti in merito.

Teorema 5.2.1 (Teorema delle dimensioni) *Data una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W, con $\dim V = n$, allora*

$$\dim V = \dim(\mathrm{Im}L) + \dim(\mathrm{Ker}L).$$

Dimostrazione. (N.B. Non stiamo facendo alcuna ipotesi su W, in ogni caso tutto dipende dal dominio!). Essendo $\mathrm{Ker}L \leq V$ e $\dim V = n$, anche $\mathrm{Ker}L$ ha dimensione finita, sia essa $k \leq n$ e sia $\{t_1, \ldots, t_k\}$ una base di $\mathrm{Ker}L$. I vettori $t_1, \ldots, t_k$ sono linearmente indipendenti anche in V e si possono allora completare in una base di V (Teorema 3.2.8): $\{t_1, \ldots, t_k, t_{k+1}, \ldots, t_n\}$. I vettori immagine: $L(t_1), \ldots, L(t_k), L(t_{k+1}), \ldots, L(t_n)$ sono un insieme di generatori di $\mathrm{Im}L$; per costruzione $L(t_1) = \ldots = L(t_k) = \mathbf{0}_W$ e, quindi, non partecipano alla "generazione". Allora i vettori $L(t_{k+1}), \ldots, L(t_n)$ sono generatori di $\mathrm{Im}L$. La proposizione sarà dimostrata se dimostreremo che $L(t_{k+1}), \ldots, L(t_n)$ sono vettori linearmente indipendenti nel qual caso $\dim(\mathrm{Im}L) = n-k$, quindi $n = \dim V = k+(n-k) = \dim(\mathrm{Ker}L)+\dim(\mathrm{Im}L)$. Sia dunque $\beta_1 L(t_{k+1}) + \cdots + \beta_{n-k} L(t_n) = \mathbf{0}_W$ una relazione di dipendenza tra i vettori $L(t_{k+1}), \ldots, L(t_n)$, con $\beta_i \in \mathbb{R}$, cioè, per la linearità di L, $L(\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n) = \mathbf{0}_W$. Questo significa che $\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n \in \mathrm{Ker}L$, ma noi avevamo scelto una base di V tale che fosse $\langle t_1, \ldots, t_k \rangle \oplus \langle t_{k+1}, \ldots, t_n \rangle = \mathrm{Ker}L \oplus \langle t_{k+1}, \ldots, t_n \rangle$, quindi tra le combinazioni lineari dei vettori linearmente indipendenti $t_{k+1}, \ldots, t_n$ non ci può essere un vettore del nucleo se non il vettore nullo. Così $\beta_1 t_{k+1} + \cdots + \beta_{n-k} t_n = \mathbf{0}_V$ e, dal momento che i vettori $t_{k+1}, \ldots, t_n$ sono linearmente indipendenti, i β_i sono nulli. Quindi i vettori $L(t_{k+1}), \ldots, L(t_n)$ sono linearmente indipendenti. **C.V.D.**

Osservazione 5.2.2 i) La dimensione dell'immagine di una applicazione lineare è sempre più piccola della dimensione del dominio o uguale ad essa.

ii) Si dice che una applicazione lineare è un *endomorfismo di V* se è un'applicazione lineare in cui dominio e codominio coincidono con V cioè: $L : V \to V$. Se un endomorfismo di V è suriettivo allora è pure iniettivo: infatti se la funzione è suriettiva ($\mathrm{Im}L = V$) allora $\dim(\mathrm{Im}L) = \dim(V)$ il che implica, per il teorema 5.2.1, $\dim(\mathrm{Ker}L) = 0$ cioè il nucleo è il sottospazio banale, quindi la funzione è iniettiva. Viceversa, se l'endomorfismo è iniettivo, $\mathrm{Ker}L = \{\mathbf{0}_V\}$ e dunque la sua dimensione è zero. Allora $\dim V = \dim L$ e quindi $\mathrm{Im}L$ ha dimensione uguale alla dimensione dello spazio di cui è sottospazio, dunque coincide con esso e la funzione è suriettiva. In conclusione, un endomorfismo di uno spazio vettoriale V è iniettivo se e solo se è suriettivo e quindi biiettivo. Si osservi che questa proprietà differenzia in modo sostanziale il comportamento delle funzioni lineari da quello delle funzioni non lineari. Si faccia, per esercizio, un esempio di una funzione (non lineare) $g : \mathbb{R} \to \mathbb{R}$ iniettiva ma non suriettiva o suriettiva ma non iniettiva.

iii) Sia $L : V \to W$ un'applicazione lineare tra spazi vettoriali di dimensione finita.

Se $\dim V = n > \dim W = m$, per il Teorema 5.2.1, $n = \dim \mathrm{Ker}(L) + \dim \mathrm{Im}(L)$, ma, essendo $\mathrm{Im}(L) \leq W$, si ha $\dim \mathrm{Im}(L) \leq m$ quindi $\dim \mathrm{Ker}(L) = n - \dim \mathrm{Im}(L) \geq n - m > 0$. Essendo $\dim \mathrm{Ker}(L) > 0$ l'applicazione lineare L non può mai essere iniettiva.

Se $\dim \mathrm{Ker}(L) + \dim \mathrm{Im}(L) = n < m$ si ha $\dim \mathrm{Im}(L) = n - \dim \mathrm{Ker}(L) \leq n < m$, quindi l'applicazione lineare L non può mai essere suriettiva.

iv) Una applicazione lineare $L : V \to W$ tra due $\mathbb{R}$-spazi vettoriali V, W di dimensione finita, si dice un **isomorfismo** se è biiettiva. Poiché la funzione è suriettiva $\mathrm{Im}L = W$, poiché la funzione è iniettiva $\dim(\mathrm{Ker}L) = 0$ e dal Teorema delle dimensioni deduciamo che $\dim V = \dim(\mathrm{Im}L) =$

dim W. Si noti che non è detto a priori che la funzione inversa di un'applicazione lineare sia lineare, ma dimostreremo in 7.2.2. che, di fatto, lo è.

Esempio 5.2.3 Sia V uno spazio vettoriale e consideriamo una sua decomposizione in somma diretta $V = V_1 \oplus V_2$ ove V_1, V_2 sono suoi sottospazi. Sappiamo allora che ogni vettore $v \in V$ si decompone in modo unico come somma $v = v_1 + v_2$ ove $v_1 \in V_1$ e $v_2 \in V_2$. Possiamo chiamare v_1 la proiezione di v su V_1 lungo V_2 e la denotiamo con $p_{V_1}^{V_2}(v) = v_1$. Si noti che essendo V_1 un sottospazio di V, si ha che $p_{V_1}^{V_2}(v) \in V$ per ogni $v \in V$. Possiamo definire l'applicazione proiezione su V_1 nella direzione di V_2

$$\begin{array}{rccc} p_{V_1}^{V_2}: & V & \longrightarrow & V \\ & v & \longmapsto & p_{V_1}^{V_2}(v) = v_1 \end{array}$$

Mostriamo che l'applicazione $p_{V_1}^{V_2}$ è lineare e quindi un endomorfismo di V. Siano $v, w \in V$. Dobbiamo mostrare che $p_{V_1}^{V_2}(v + w) = p_{V_1}^{V_2}(v) + p_{V_1}^{V_2}(w)$. Essendo la somma diretta possiamo scrivere in modo unico $v = v_1 + v_2$ e $w = w_1 + w_2$, con $v_1, w_1 \in V_1$ e $v_2, w_2 \in V_2$. Ne segue che $v + w = (v_1 + v_2) + (w_1 + w_2)$ e, per la commutatività e l'associatività della somma, si ha $v + w = (v_1 + w_1) + (v_2 + w_2)$ con $v_1 + w_1 \in V_1$ e $v_2 + w_2 \in V_2$ (essendo V_1 e V_2 sottospazi di V). Il fatto che la somma di V_1 e V_2 sia diretta ci assicura che questa è l'unica decomposizione di $v + w$. Di conseguenza $p_{V_1}^{V_2}(v + w) = v_1 + w_1 = p_{V_1}^{V_2}(v) + p_{V_1}^{V_2}(w)$ che è quello che si voleva provare. Resta da verificare che, presi $\lambda \in \mathbb{R}$ e $v \in V$, si ha $p_{V_1}^{V_2}(\lambda v) = \lambda p_{V_1}^{V_2}(v)$. Ma se $v = v_1 + v_2$ con $v_1 \in V_1$ e $v_2 \in V_2$ allora $\lambda v = \lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$ con $\lambda v_1 \in V_1$ e $\lambda v_2 \in V_2$ (perché V_1 e V_2 sono sottospazi vettoriali di V). Come prima si ha che $p_{V_1}^{V_2}(\lambda v) = \lambda v_1 = \lambda p_{V_1}^{V_2}(v)$ e la funzione $p_{V_1}^{V_2}$ è dunque lineare. La funzione $p_{V_1}^{V_2}$ si dirà **la proiezione** su V_1 lungo V_2.

Analogamente si definisce simmetria di asse V_1 di direzione V_2 l'applicazione lineare:

$$\begin{array}{rccc} \sigma_{V_1}^{V_2}: & V & \longrightarrow & V \\ & v & \longmapsto & \sigma_{V_1}^{V_2}(v) = v_1 - v_2. \end{array}$$

Tale applicazione è un isomorfismo e la sua inversa è la simmetria stessa.

5.3 Applicazioni lineari, basi e matrici

Abbiamo visto che, dato uno spazio vettoriale V e fissata una sua base $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$, ogni vettore $v \in V$ è univocamente individuato dalla n-upla delle sue coordinate rispetto alla base scelta: se $v = \alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n$, allora v può essere indicato con $(\alpha_1, \alpha_2, \ldots, \alpha_n)_{\mathbf{v}}$. Questa n-upla è una specie di numero di maglia che diamo ad ogni vettore/giocatore: il colore della maglia ci dice di quale squadra si tratta e il numero sulla maglia individua univocamente il giocatore. Lo stesso numero su maglie di colore diverso individua giocatori diversi così come la stessa n-upla di numeri reali rispetto a basi diverse individua vettori diversi. Ad esempio, se in $\mathbb{R}^2$ scegliamo le due basi $\mathbf{v} = \{v_1 = (1, 2), v_2 = (1, 1)\}$ e $\mathbf{w} = \{w_1 = (1, 0), w_2 = (1, 5)\}$, il vettore di coordinate $(1, 1)_{\mathbf{v}}$ nella base $\mathbf{v}$ è il vettore $1v_1 + 1v_2 = 1(1, 2) + 1(1, 1) = (2, 3)$, mentre $(1, 1)_{\mathbf{w}}$ è il vettore $1w_1 + 1w_2 = 1(1, 0) + 1(1, 5) = (2, 5)$, stesso numero ma su maglie diverse: giocatori diversi. In ogni caso, se lo spazio ha dimensione n, la scelta di una base ci assicura di poter scrivere ogni vettore come una $n - upla$ di numeri reali.

Come possiamo costruire una applicazione lineare nel modo più semplice possibile? Sia $L : V \to W$ una applicazione lineare tra due spazi di dimensione finita e siano $\dim V = n$ e $\dim W = m$; scegliamo poi una base per ciascuno spazio: $\mathbf{v} = \{v_1, v_2, \ldots, v_n\}$ e $\mathbf{w} = \{w_1, w_2, \ldots, w_m\}$. Abbiamo visto in 5.1.5.iv) (e lo ricordiamo) che conoscere i valori di L sui vettori di una base significa conoscere l'applicazione lineare interamente. Infatti, supponiamo di sapere chi siano $L(v_1), L(v_2), \ldots, L(v_n) \in W$ e di voler calcolare l'immagine di ogni vettore $v \in V$. Poiché i vettori $v_1, v_2, \ldots, v_n$ sono una base di V, v si scrive in modo unico come loro combinazione lineare: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e, applicando L, si ha: $L(v) = L(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda_1 L(v_1) + \lambda_2 L(v_2) + \cdots + \lambda_n L(v_n)$. Dunque conoscendo $L(v_1), L(v_2), \ldots, L(v_n)$ e sfruttando l'unicità della scrittura di v nella base $v_1, v_2, \ldots, v_n$, si conosce la funzione lineare L completamente. Ora, poiché abbiamo fissato una base di W, ogni vettore $L(v_1), \ldots, L(v_n)$ si può scrivere in modo unico come combinazione lineare dei vettori di questa base:

$$
\begin{array}{rcl}
L(v_1) & = & a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\
L(v_2) & = & a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\
\vdots & & \\
L(v_n) & = & a_{1n}w_1 + a_{2n}w_2 + \ldots + a_{mn}w_m.
\end{array}
$$

I termini a_{ij} sono univocamente determinati dalle scelte da noi fatte delle basi di V e W (si noti che fissare una base significa fissare anche l'ordine dei suoi elementi). Possiamo allora costruire la seguente matrice:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Tale matrice in $\mathcal{M}_{m,n}(\mathbb{R})$ dipende dalla scelta delle due basi $\{v_1, v_2, \dots, v_n\}$ e $\{w_1, w_2, \dots, w_m\}$, ma caratterizza completamente la nostra applicazione lineare L. La matrice A si dice **la matrice associata alla applicazione lineare L rispetto alle basi $\mathbf{v}$ e $\mathbf{w}$**. Osserviamo che, una volta fissate le basi $\mathbf{v}$ e $\mathbf{w}$, le colonne di A rappresentano le coordinate rispetto alla base $\mathbf{w}$ dei vettori $L(v_1), \dots, L(v_n)$. A questo punto l'immagine di un vettore $v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n \in V$ si può esprimere in coordinate rispetto alla base $\{w_1, w_2, \dots, w_m\}$, come segue:

$$\begin{aligned} L(v) &= \lambda_1 L(v_1) + \lambda_2 L(v_2) + \dots + \lambda_n L(v_n) = \\ &= \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ \dots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ \dots \\ a_{m2} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ \dots \\ a_{mn} \end{pmatrix} = \\ &= \begin{pmatrix} \lambda_1 a_{11} + \lambda_2 a_{12} + \dots + \lambda_n a_{1n} \\ \lambda_1 a_{21} + \lambda_2 a_{22} + \dots + \lambda_n a_{2n} \\ \dots \quad \dots \quad \dots \\ \dots \quad \dots \quad \dots \\ \lambda_1 a_{m1} + \lambda_2 a_{m2} + \dots + \lambda_n a_{mn} \end{pmatrix}. \end{aligned}$$

(Si noti che le coordinate dei vettori $L(v_i)$ sono indicate come vettori colonna).

Questo ci permette di affermare che l'immagine del vettore $v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$ è data in coordinate dal prodotto

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \dots \\ \dots \\ \lambda_n \end{pmatrix}.$$

Con questa scrittura (o piuttosto, come vedremo in seguito, con il risultato di questa scrittura) si intende una matrice a una colonna e m righe: l'entrata alla riga k-esima di tale matrice è il "prodotto" della riga k della matrice A per la colonna data dalle coordinate del vettore di cui vogliamo conoscere l'immagine. Questo "prodotto" è per definizione la somma di $a_{k1}\lambda_1$, $a_{k2}\lambda_2$, e così via, fino ad ottenere $\sum_{i=1}^{n} a_{ki}\lambda_i$. (Questo prodotto, che sarà detto prodotto"righe per colonne", verrà introdotto e studiato nel paragrafo 7.1).

Esempio 5.3.1 *i*) Consideriamo la base $\mathbf{v} = \{v_1 = (1,1,0),\ v_2 = (0,1,0),\ v_3 = (2,0,1)\}$ di $\mathbb{R}^3$ e la base $\mathbf{w} = \{w_1 = (1,0), w_2 = (1,1)\}$ di $\mathbb{R}^2$ e sia $L : \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare definita da $L(v_1) = (1,-1), L(v_2) = (0,-1), L(v_3) = (2,1)$. Chi è l'immagine mediante L di un vettore di $\mathbb{R}^3$, ad esempio $v = (1,1,1)$? Scriviamo prima di tutto v rispetto alla base v_1, v_2, v_3. Si ha $v = (1,1,1) = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \lambda_1(1,1,0) + \lambda_2(0,1,0) + \lambda_3(2,0,1) = (\lambda_1 + 2\lambda_3, \lambda_1 + \lambda_2, \lambda_3)$, cioè: $\lambda_3 = 1$, $\lambda_2 = 2$, $\lambda_1 = -1$. Pertanto, per la linearità di L, abbiamo: $L(1,1,1) = -1L(v_1) + 2L(v_2) + 1L(v_3) = -(1,-1) + 2(0,-1) + (2,1) = (1,0)$. Più in generale, per ogni vettore di coordinate $(\lambda_1, \lambda_2, \lambda_3)$ nella base v_1, v_2, v_3, si ha $L(\lambda_1, \lambda_2, \lambda_3) = \lambda_1(1,-1) + \lambda_2(0,-1) + \lambda_3(2,1) = (\lambda_1 + 2\lambda_3, -\lambda_1 + \lambda_3 - \lambda_2)$. Cerchiamo la matrice associata a L rispetto alle basi fissate. Quale forma avrà tale matrice? Il numero di colonne è pari alla dimensione del dominio e il numero di righe pari alla dimensione del codominio. Quindi stiamo cercando una matrice in $\mathcal{M}_{2,3}(\mathbb{R})$. Le sue colonne sono date dalle coordinate nella base w_1, w_2 delle immagini dei vettori della base fissata nel dominio: $L(v_1) = (1,-1) = 2(1,0) - (1,1)$, $L(v_2) = (0,-1) = (1,0) - (1,1)$, $L(v_3) = (2,1) = (1,0) + (1,1)$. Otteniamo così la matrice

$$\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

L'immagine del vettore di $\mathbb{R}^3$ che nella base $\{v_1, v_2, v_3\}$ ha coordinate $(1,2,3)$ è il vettore

$$\begin{pmatrix} 2 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$$

espresso in coordinate rispetto alla base w_1 e w_2, si tratta cioè del vettore $7(1,0) + 0(1,1) = (7,0)$.

ii) Cosa succede se nell'esempio precedente cambiamo le basi degli spazi V e W? Prendiamo ad esempio la base $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$ per il dominio e la base $e' = \{e'_1 = (1,0),\ e'_2 = (0,1)\}$ per il codominio. Allora per descrivere la matrice associata a L rispetto alle nuove basi dobbiamo calcolare le coordinate dei vettori $L(e_1), L(e_2), L(e_3)$ nella base e'_1, e'_2. Ora $e_1 = v_1 - v_2$ e quindi $L(e_1) = L(v_1 - v_2) = L(v_1) - L(v_2) = (1,-1) - (0,-1) = (1,0)$; $e_2 = v_2$ e dunque $L(v_2) = L(e_2) = (0,-1)$, infine $e_3 = -2v_1 + 2v_2 + v_3$ quindi $L(e_3) = L(-2v_1 + 2v_2 + v_3) = -2L(v_1) + 2L(v_2) + L(v_3) = -2(1,-1) + 2(0,-1) + (2,1) = (0,1)$. Tali immagini sono espresse nella base $\{e'_1, e'_2\}$ come segue: $L(e_1) = 1e'_1 + 0e'_2 = (1,0)_{e'}$, $L(e_2) = 0e'_1 + (-1)e'_2 = (0,-1)_{e'}$, $L(e_3) = 0e'_1 + 1e'_2 = (0,1)_{e'}$. Otteniamo pertanto la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Osservazione 5.3.2 Sia V uno spazio vettoriale di dimensione n su $\mathbb{R}$. Fissare una base di V significa costruire un isomorfismo tra V e $\mathbb{R}^n$. Infatti sia $v_1, \ldots, v_n$ una base di V. Allora ogni vettore v di V si scrive in modo unico come combinazione lineare di $v_1, \ldots, v_n$: $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Definiamo la funzione

$$\varphi : V \longrightarrow \mathbb{R}^n$$

che associa ad ogni vettore $v \in V$ le sue coordinate nella base $v_1, \ldots, v_n$: $(\lambda_1, \lambda_2, \ldots, \lambda_n)$. La funzione φ è lineare: se $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ e $v' = \lambda'_1 v_1 + \lambda'_2 v_2 + \cdots + \lambda'_n v_n$ sono due vettori di V allora $v + v' = (\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) + (\lambda'_1 v_1 + \lambda'_2 v_2 + \cdots + \lambda'_n v_n) = (\lambda_1 + \lambda'_1) v_1 + (\lambda_2 + \lambda'_2) v_2 + \cdots + (\lambda_n + \lambda'_n) v_n$, cosicché $\varphi(v + v') = (\lambda_1 + \lambda'_1, \lambda_2 + \lambda'_2, \ldots, \lambda_n + \lambda'_n) = (\lambda_1, \lambda_2, \ldots, \lambda_n) + (\lambda'_1, \lambda'_2, \ldots, \lambda'_n) = \varphi(v) + \varphi(v')$. Ancora, se $\lambda \in \mathbb{R}$ e $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$, allora $\lambda v = \lambda(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = \lambda\lambda_1 v_1 + \lambda\lambda_2 v_2 + \cdots + \lambda\lambda_n v_n$, dunque $\varphi(\lambda v) = (\lambda\lambda_1, \lambda\lambda_2, \ldots, \lambda\lambda_n) = \lambda\varphi(v)$. Quindi φ è lineare. Ed è in particolare una funzione lineare tra due spazi della stessa dimensione. Per il Teorema delle dimensioni se φ è iniettiva allora è pure suriettiva. Del resto φ è iniettiva perché se $\varphi(v) = \mathbf{0}_{\mathbb{R}^n} = (0, 0, \ldots, 0)$ significa che le coordinate del vettore v sono nulle cioè: $v = 0v_1 + 0v_2 + \ldots + 0v_n = \mathbf{0}_V$. La funzione è iniettiva e quindi un isomorfismo. In particolare questo significa che ragionare sui vettori di V equivale a ragionare sui vettori pensati attraverso le loro coordinate rispetto ad una base fissata o, equivalentemente, che ogni spazio vettoriale reale di dimensione n è “sostanzialmente” $\mathbb{R}^n$.

Da quanto detto sopra dovrebbe risultare chiaro come poter calcolare la dimensione dell'immagine di una applicazione lineare quando è data la matrice ad essa associata rispetto a basi fissate. Spieghiamolo meglio: siano allora $L : V \to W$ un'applicazione lineare, $\mathbf{v} = \{v_1, \ldots, v_n\}$ una base di V e

$\mathbf{w} = \{w_1, w_2, \ldots, w_m\}$ una base di W. Le coordinate rispetto a $\mathbf{w}$ dei vettori $L(v_1)$, $L(v_2), \ldots, L(v_n)$ sono le colonne della matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ associata ad L. Ora dobbiamo calcolare $\dim(\mathrm{Im}L) = \dim\langle L(v_1), L(v_2), \ldots, L(v_n)\rangle$ (vedi 5.2), cioè la dimensione del sottospazio vettoriale generato da $L(v_1)$, $L(v_2)$, ..., $L(v_n)$. Questo è equivalente a determinare il massimo numero di vettori linearmente indipendenti tra $L(v_1)$, $L(v_2)$, ..., $L(v_n)$ cioè il massimo numero di colonne di A linearmente indipendenti, pensando ogni colonna di A come un vettore di coordinate rispetto alla base $\{w_1, w_2, \ldots, w_m\}$, e quindi come un vettore di $\mathbb{R}^m$. Diamo dunque la seguente definizione:

Definizione 5.3.3 *Data una matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$, diremo rango colonne di A, e lo indicheremo con* $\mathrm{rg}A$*, il massimo numero di colonne linearmente indipendenti di A (pensate come vettori di $\mathbb{R}^m$).*

In conclusione, se $L : V \to W$ è una applicazione lineare tra due $\mathbb{R}$-spazi vettoriali, con $\dim V = n$ e $\dim W = m$, e se la matrice associata a tale applicazione lineare rispetto ad una base fissata del dominio ed una del codominio è $A \in \mathcal{M}_{m,n}(\mathbb{R})$, allora $\mathrm{rg}A = \dim(\mathrm{Im}L)$. Inoltre per ogni possibile scelta di basi di V, W tutte le matrici associate a L appartengono all'insieme $\mathcal{M}_{m,n}(\mathbb{R})$ e tutte hanno rango colonne pari a $\dim(\mathrm{Im}L)$. Quindi il rango colonne dipende solo dall'applicazione lineare L e non dalle basi scelte nel dominio e nel codominio.

5.4 Esercizi svolti

Esercizio 5.4.1 Stabilire quali delle seguenti applicazioni sono lineari:

i) $f_1 : \mathbb{R} \longrightarrow \mathbb{R}$, $f_1(x) = x + 3$.

ii) $f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $f_2(x, y) = (x^2, y)$.

iii) $f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f_3(x, y) = 2x + 3y$.

iv) $f_4 : \mathcal{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}^2$, $f_4 \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a + c, b + d)$.

v) $f_5 : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathcal{M}_n(\mathbb{R})$, $f_5(A) = 2A$.

Svolgimento.

i) L'immagine mediante un'applicazione lineare del vettore nullo del dominio è sempre il vettore nullo del codominio, ma $f_1(0) = 3$ quindi f_1 non è un'applicazione lineare.

ii) L'applicazione f_2 non è lineare dal momento che $f_2((1,0) + (-1,0)) = f_2(0,0) = (0,0) \neq f_2(1,0) + f_2(-1,0) = (1,0) + (1,0) = (2,0)$.

iii) Per verificare che un'applicazione $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ è un'applicazione lineare occorre verificare le due seguenti condizioni:

   1) per ogni coppia di vettori v e w in $\mathbb{R}^2$: $f(v+w) = f(v) + f(w)$;
   2) per ogni $v \in \mathbb{R}^2$ e per ogni $\alpha \in \mathbb{R}$ $f(\alpha v) = \alpha f(v)$.

   Consideriamo dunque l'applicazione f_3 e siano $v = (a,b)$ e $w = (c,d)$ due elementi di $\mathbb{R}^2$. Abbiamo:

   $f_3(v+w) = f_3(a+c, b+d) = 2(a+c)+3(b+d) = (2a+3b)+(2c+3d) = f_3(v) + f_3(w)$, quindi la proprietà 1) è verificata.

   Analogamente, preso α in $\mathbb{R}$,

$$f_3(\alpha v) = f_3(\alpha a, \alpha b) = 2\alpha a + 3\alpha b = \alpha(2a+3b) = \alpha f_3(v).$$

   Possiamo concludere che l'applicazione f_3 è lineare.

Analogamente si procede per dimostrare che le applicazioni f_4 e f_5 sono lineari.

Esercizio 5.4.2 Sia $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione definita da:

$$f(x,y,z) = (x+y, x+y, z).$$

i) Scrivere la matrice associata a f rispetto alla base canonica nel dominio e nel codominio.

ii) Determinare $\mathrm{Ker} f$ e $\mathrm{Im} f$.

iii) Scrivere la matrice associata a f rispetto alla base canonica nel dominio e alla base $\{v_1 = (1,1,1), v_2 = (1,1,0), v_3 = (1,-1,0)\}$ nel codominio.

iv) Scrivere la matrice associata a f rispetto alla base $\{v_1, v_2, v_3\}$ nel dominio e alla base canonica nel codominio.

Svolgimento.

i) La matrice associata a f rispetto alla base canonica nel dominio e nel codominio è la matrice che ha sulle colonne le coordinate rispetto alla base canonica di $\mathbb{R}^3$ delle immagini mediante f dei vettori della stessa base. Calcoliamo dunque:

$$f(1,0,0) = (1,1,0) = 1(1,0,0) + 1(0,1,0) + 0(0,0,1)$$

$$f(0,1,0) = (1,1,0) = 1(1,0,0) + 1(0,1,0) + 0(0,0,1)$$

$$f(0,0,1) = (0,0,1) = 0(1,0,0) + 0(0,1,0) + 1(0,0,1).$$

Dunque la matrice associata a f rispetto alla base canonica nel dominio e nel codominio è:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

ii) Sappiamo che Imf è generata dai vettori $f(1,0,0)$, $f(0,1,0)$ e $f(0,0,1)$. Dunque Im$f = \langle (1,1,0), (0,0,1) \rangle$. Usando il Teorema delle dimensioni 5.2.1 otteniamo che Ker(f) ha dimensione 1. Del resto, dal momento che $f(1,0,0) = (1,1,0) = f(0,1,0)$, per la linearità di f il vettore $(1,0,0) - (0,1,0) = (1,-1,0)$ ha come immagine mediante f il vettore nullo: $f((1,0,0) - (0,1,0)) = f(1,0,0) - f(0,1,0) = 0_{\mathbb{R}^3}$. Dunque Ker$(f) = \langle (1,-1,0) \rangle$.

iii) La matrice richiesta ha sulle colonne le coordinate nella base $\{v_1, v_2, v_3\}$ delle immagini dei vettori della base canonica. Abbiamo già determinato le immagini, tramite f, dei vettori della base canonica. Si tratta ora di esprimere queste immagini in coordinate rispetto alla base $\{v_1, v_2, v_3\}$. Abbiamo:

$$\begin{array}{rcccl} f(1,0,0) & = & (1,1,0) & = & 0v_1 + 1v_2 + 0v_3 \\ f(0,1,0) & = & (1,1,0) & = & 0v_1 + 1v_2 + 0v_3 \\ f(0,0,1) & = & (0,0,1) & = & 1v_1 - 1v_2 + 0v_3. \end{array}$$

La matrice richiesta è dunque:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

iv) In questo caso la base fissata nel dominio è la base $\{v_1, v_2, v_3\}$. Calcoliamo allora le immagini dei vettori v_1, v_2, v_3 e determiniamo le coordinate dei vettori trovati rispetto alla base canonica:

$$\begin{array}{rclcl} f(v_1) & = & (2,2,1) & = & 2(1,0,0)+2(0,1,0)+1(0,0,1) \\ f(v_2) & = & (2,2,0) & = & 2(1,0,0)+2(0,1,0)+0(0,0,1) \\ f(v_3) & = & (0,0,0). & & \end{array}$$

La matrice richiesta è dunque:

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Esercizio 5.4.3 Sia $D : \mathbb{R}^{\leq 3}[x] \longrightarrow \mathbb{R}^{\leq 3}[x]$ l'applicazione derivata (rispetto alla variabile x). Determinare Ker$D$, ImD e la matrice associata a D rispetto alla base $\mathcal{B} = \{1, x, x^2, x^3\}$ di $\mathbb{R}^{\leq 3}[x]$.

Svolgimento. Ricordiamo la definizione di derivata di un polinomio in una variabile (di grado ≤ 3):

$$D(a_0 + a_1 x + a_2 x^2 + a_3 x^3) = a_1 + 2a_2 x + 3a_3 x^2.$$

Si verifica immediatamente, usando questa definizione, che la derivata è un'applicazione lineare. Per definizione di nucleo di un'applicazione lineare,

$$\begin{array}{rcl} \mathrm{Ker}(D) & = & \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid D(p(x)) = 0\} = \\ & = & \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_1 + 2a_2x + 3a_3x^2 = 0\}. \end{array}$$

Dunque

$$\begin{array}{rcl} \mathrm{Ker}(D) & = & \{p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_1 = a_2 = a_3 = 0\} = \\ & = & \mathbb{R}. \end{array}$$

Usando 5.2.1 deduciamo immediatamente che l'immagine di D ha dimensione 3. Del resto

$$\begin{array}{rcl} \mathrm{Im}D & = & \langle D(1), D(x), D(x^2), D(x^3)\rangle = \\ & = & \langle 1, 2x, 3x^2\rangle. \end{array}$$

La matrice associata a D rispetto alla base $\mathcal{B}$ è:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Esercizio 5.4.4 Possiamo, nelle ipotesi precedenti, considerare l'applicazione composta $D \circ D = D^2$, cioè la derivata seconda nella variabile x, come applicazione di $\mathbb{R}^{\leq 3}[x]$ in se stesso. Si verifichi che D^2 è una applicazione lineare. Definiamo allora l'applicazione che ad ogni polinomio di $\mathbb{R}^{\leq 3}[x]$ associa la sua derivata seconda meno la sua derivata prima: $P \longmapsto D^2(P) - D(P)$. Si verifichi che anche questo è un endomorfismo di $\mathbb{R}^{\leq 3}[x]$. In particolare si studi il suo nucleo. Indicando, secondo la consuetudine, la derivata prima e la derivata seconda di P rispettivamente con P' e P'', il nucleo dell'applicazione costruita è allora dato dall'insieme di polinomi P che soddisfano la relazione

$$P'' - P' = 0.$$

L'equazione trovata è detta equazione differenziale.

Esercizio 5.4.5 Tra le applicazioni lineari dell'esercizio 5.4.1 si determinino quelle iniettive e quelle suriettive.

Svolgimento. Le funzioni f_1 e f_2 non sono lineari.
La funzione f_3 non è iniettiva:$f_3(-3,2) = 0$, dunque $(-3,2)$ è un vettore non nullo del nucleo di f_3. D'altra parte dal teorema delle dimensioni si può concludere che $\dim \mathbb{R}^2 = \dim \operatorname{Ker}(f_3) + \dim \operatorname{Im}(f_3)$ da cui $\dim \operatorname{Ker}(f_3) = 2 - \dim \operatorname{Im}(f_3) \geq 1$, quindi non esistono applicazioni lineari da $\mathbb{R}^2$ a $\mathbb{R}$ iniettive (vedi anche 5.2.2.iii). Nel nostro caso f_3 è suriettiva, infatti $\operatorname{Im} f_3$ contiene il vettore $f_3(1,0) = 2$ dunque ha dimensione almeno 1. Del resto $\dim(\mathbb{R}) = 1$, quindi $\operatorname{Im} f_3 = \mathbb{R}$.

Nello stesso modo f_4 è un'applicazione lineare suriettiva ma non iniettiva.

La funzione f_5 è un endomorfismo, dunque essa è iniettiva se e solo se è suriettiva. Del resto f_5 è ovviamente suriettiva e quindi biiettiva.

Esercizio 5.4.6 Sia $L : V \longrightarrow W$ un'applicazione lineare e siano $\{v_1, \ldots, v_n\}$ vettori linearmente indipendenti di V. Mostrare che se L è iniettiva allora i vettori $\{L(v_1), \ldots, L(v_n)\}$ di W sono linearmente indipendenti.

Svolgimento. Consideriamo una combinazione lineare dei vettori $L(v_1)$, $\ldots, L(v_n)$ e supponiamo che essa sia uguale al vettore nullo di W:

$$\alpha_1 L(v_1) +_W \cdots +_W \alpha_n L(v_n) = \mathbf{0}_W.$$

Per la linearità di L si ha:

$$0_W = \alpha_1 L(v_1) +_W \cdots +_W \alpha_n L(v_n) = L(\alpha_1 v_1 +_V \cdots +_V \alpha_n v_n).$$

Essendo L iniettiva per ipotesi, l'unico vettore di V che ha come immagine 0_W è il vettore nullo di V, quindi:

$$\alpha_1 v_1 +_V \cdots +_V \alpha_n v_n = \mathbf{0}_V$$

e questo implica $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ dal momento che i vettori $v_1, \ldots, v_n$ sono linearmente indipendenti. L'esercizio è concluso.

Esercizio 5.4.7 i) È possibile costruire una applicazione lineare iniettiva da $\mathbb{R}^3$ in $\mathbb{R}^2$? E un'applicazione suriettiva? In caso affermativo si facciano degli esempi.

ii) È possibile costruire una applicazione lineare iniettiva da $\mathbb{R}^2$ in $\mathbb{R}^3$? E un'applicazione suriettiva? In caso affermativo si facciano degli esempi.

Svolgimento.

i) Dall'osservazione 5.2.2iii), sappiamo già che non vi sono applicazioni lineari iniettive fra $\mathbb{R}^3$ e $\mathbb{R}^2$. Come ulteriore verifica, notiamo che nell'esercizio precedente abbiamo mostrato che un'applicazione lineare iniettiva manda vettori linearmente indipendenti in vettori linearmente indipendenti dunque non esiste un'applicazione lineare iniettiva da uno spazio di dimensione n in uno spazio di dimensione m se $m < n$.

   Al contrario è certamente possibile costruire un'applicazione lineare suriettiva da $\mathbb{R}^3$ in $\mathbb{R}^2$, ad esempio l'applicazione lineare f definita da: $f(1, 0, 0) = (1, 0)$, $f(0, 1, 0) = (0, 1)$, $f(0, 0, 1) = (1, 1)$ è suriettiva dal momento che $\mathrm{Im} f = \langle (1, 0), (0, 1) \rangle = \mathbb{R}^2$.

ii) Sia $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da: $g(1, 0) = (1, 0, 0)$, $g(0, 1) = (0, 1, 0)$. L'applicazione lineare g è iniettiva, infatti $\mathrm{Im} g = \langle g(1, 0), g(0, 1) \rangle = \langle (1, 0, 0), (0, 1, 0) \rangle$ ha dimensione 2, di conseguenza il nucleo di g ha dimensione 0, cioè è banale.

   D'altra parte, sempre per 5.2.2iii), l'immagine di un'applicazione lineare ha dimensione minore della dimensione del dominio o uguale ad essa. Dunque non è possibile costruire un'applicazione lineare suriettiva da $\mathbb{R}^2$ in $\mathbb{R}^3$.

Esercizio 5.4.8 i) Esiste un'applicazione lineare $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ tale che $g(1, 0, 0, 0) = (3, 4)$, $g(2, 3, 4, 0) = (1, 2)$, e $g(0, 3, 4, 0) = (6, 5)$? In caso affermativo se ne faccia un esempio.

ii) Esiste un'applicazione lineare $h : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ tale che $h(1,0,0,0) = (3,4)$, $h(2,3,4,0) = (1,2)$, $h(0,0,3,2) = (2,8)$? In caso affermativo le si descrivano tutte.

Svolgimento.

i) Osserviamo innanzitutto che $(0,3,4,0) = (2,3,4,0) - 2(1,0,0,0)$. Pertanto, se esistesse una funzione lineare g come richiesta, si avrebbe: $g(0,3,4,0) = g(2,3,4,0) - 2g(1,0,0,0)$, cioè $(6,5) = (1,2) - 2(3,4)$, ma questo non è possibile.

ii) Osserviamo che i vettori $(1,0,0,0)$, $(2,3,4,0)$, $(0,0,3,2)$ sono linearmente indipendenti (si verifichi!) dunque è possibile definire un'applicazione lineare fissando liberamente le loro immagini. In particolare esiste sicuramente un'applicazione lineare h come richiesta.

Per definire un'applicazione lineare è sufficiente definire le immagini dei vettori di una base del dominio. Nel nostro caso possiamo aggiungere all'insieme $(1,0,0,0)$, $(2,3,4,0)$, $(0,0,3,2)$ il vettore $(0,0,0,1)$ per ottenere una base di $\mathbb{R}^4$. Allora una qualsiasi applicazione h soddisfacente le ipotesi dell'esercizio sarà definita da $h(1,0,0,0) = (3,4)$, $h(2,3,4,0) = (1,2)$, $h(0,0,3,2) = (2,8)$ $h(0,0,0,1) = (a,b)$ al variare di a e b in $\mathbb{R}$. Notiamo, in particolare, che esistono infinite applicazioni lineari soddisfacenti le ipotesi.

Esercizio 5.4.9 Sia L l'endomorfismo di $\mathbb{R}^3$ che rispetto alla base $\mathcal{B} = \{(2,1,2), (3,1,1), (1,0,1)\}$ ha come matrice

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Determinare l'immagine mediante L del vettore $w = (1,1,1)$. Determinare, inoltre, nucleo e immagine di L.

Svolgimento. La matrice A ha sulle colonne le coordinate nella base $\mathcal{B}$ delle immagini dei vettori della base $\mathcal{B}$ stessa. Pertanto, se (a,b,c) sono le coordinate di un vettore v nella base $\mathcal{B}$, $L(v) = A \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ saranno le coordinate nella base $\mathcal{B}$ del vettore $L(v)$.

Dobbiamo dunque, innanzitutto, calcolare le coordinate del vettore $(1,1,1)$ nella base $\mathcal{B}$. Si verifica facilmente che $(1,1,1) = 1(2,1,2) + 0(3,1,1) - 1(1,0,1)$, cioè $(1,1,1) = (1,0,-1)_{\mathcal{B}}$. Pertanto

$$L(w) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$$

cioè $L(w) = (2,-1,0)_{\mathcal{B}} = 2(2,1,2) - (3,1,1) = (1,1,3)$.

Chi è il nucleo dell'applicazione L? È, per definizione, l'insieme dei vettori $(x_1,x_2,x_3)_{\mathcal{B}}$ tali che sia

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè i vettori $(x_1,x_2,x_3)_{\mathcal{B}}$ tali che $(2x_1+x_2, 3x_2+x_3, x_1+x_2+x_3) = (0,0,0)$, i.e., $x_1 = x_2 = x_3 = 0$. Dunque $\mathrm{Ker}L = \mathbf{0}_{\mathbb{R}^3}$ e la funzione L è iniettiva. Di conseguenza L è suriettiva (L è un endomorfismo) pertanto $ImL = \mathbb{R}^3$ e quindi il rango colonne di A è $\mathrm{rg}A = 3$.

5.5 Esercizi proposti

Esercizio 5.5.1 Stabilire quali delle seguenti applicazioni sono lineari:

1. $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x,y) = (x-y+2, x+y)$;
2. $g : \mathbb{R}^3 \to \mathbb{R}^2$, $g(x,y,z) = (x^2-y^2, x+y)$;
3. $h : \mathbb{R}^2 \to \mathcal{M}_2(\mathbb{R})$, $h(x,y) = \begin{pmatrix} x-y & 0 \\ 0 & 2x-3y \end{pmatrix}$.

Esercizio 5.5.2 Si consideri, al variare di $k \in \mathbb{R}$, l'applicazione $f_k : \mathbb{R}^2 \to \mathbb{R}^3$, $f_k(x,y) = (x-y, ky, ky)$. Stabilire per quali valori di k l'applicazione f_k è lineare. Per i valori di k trovati:

1. Scrivere la matrice associata ad f_k rispetto alla base $\mathcal{B} = \{(1,1),(1,2)\}$ di $\mathbb{R}^2$ ed alla base canonica di $\mathbb{R}^3$.
2. Determinare nucleo e immagine di f_k.

3. Determinare i valori di k tali che il vettore $(1, 0, 0)$ appartenga all'immagine di f_k.

Esercizio 5.5.3 Costruire una applicazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ non nulla tale che:

1. f non sia suriettiva;
2. il nucleo di f contenga i vettori $(2, 2, 2)$, $(1, 1, -1)$.

Scrivere la matrice associata all'applicazione f costruita, rispetto alle basi canoniche di $\mathbb{R}^3$ ed $\mathbb{R}^2$. L'applicazione f richiesta è unica?

Esercizio 5.5.4 Si consideri l'applicazione lineare $L : \mathbb{R}^{\leq 2}[x] \rightarrow \mathbb{R}^3$, $L(a + bx + cx^2) = (a + b, a + c, b - c)$.

1. Dopo aver fissato una base di $\mathbb{R}^{\leq 2}[x]$ ed una base di $\mathbb{R}^3$, determinare la matrice associata ad L rispetto a tali basi.
2. Determinare nucleo ed immagine di L. L'applicazione L è iniettiva? È suriettiva?

Esercizio 5.5.5 Sia $f : \mathcal{M}_2(\mathbb{R}) \rightarrow \mathbb{R}^2$ l'applicazione così definita:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (b - c, b + c).$$

1. Mostrare che l'applicazione f è lineare.
2. Scrivere la matrice associata ad f rispetto alle basi canoniche di $\mathcal{M}_2(\mathbb{R})$ e $\mathbb{R}^2$.
3. Determinare nucleo e immagine di f.
4. Determinare la controimmagine mediante f del sottospazio vettoriale $S = \langle (0, 1) \rangle$ di $\mathbb{R}^2$.

Lezione 6

Sistemi lineari

6.1 Applicazioni lineari vs matrici

Definizione 6.1.1 *Dati V e W spazi vettoriali reali, indichiamo con $\mathcal{L}in(V, W)$ l'insieme di tutte le applicazioni lineari tra V e W.*

Vogliamo mostrare che, dati V e W spazi vettoriali di dimensione finita rispettivamente n_1 e n_2, l'insieme $\mathcal{L}in(V, W)$ è uno spazio vettoriale di dimensione finita, e che ogni scelta di una base di V e di una di W individua un isomorfismo tra $\mathcal{L}in(V, W)$ e $\mathcal{M}_{n_2,n_1}(\mathbb{R})$. In altre parole, una volta fissate una base del dominio ed una del codominio, parlare di applicazioni lineari tra spazi vettoriali sarà del tutto equivalente a parlare di matrici. Quest'identificazione avrà delle conseguenze importanti.

Definiamo prima di tutto la struttura di spazio vettoriale su $\mathcal{L}in(V, W)$. La somma "$+_{\mathcal{L}}$" di due applicazioni lineari è definita come segue: se $f, g \in \mathcal{L}in(V, W)$, $f +_{\mathcal{L}} g$ è l'applicazione lineare tale che $(f +_{\mathcal{L}} g)(v) = f(v) +_W g(v)$ per $v \in V$. Verifichiamo che $f +_{\mathcal{L}} g$ è un'applicazione lineare. Per ogni coppia di elementi $v_1, v_2 \in V$ si ha:

$$\begin{aligned}
(f +_{\mathcal{L}} g)(v_1 +_V v_2) &= f(v_1 +_V v_2) +_W g(v_1 +_V v_2) = \\
&= f(v_1) +_W f(v_2) +_W g(v_1) +_W g(v_2) = \\
&= (f(v_1) +_W g(v_1)) +_W (f(v_2) +_W g(v_2)) = \\
&= (f +_{\mathcal{L}} g)(v_1) +_W (f +_{\mathcal{L}} g)(v_2).
\end{aligned}$$

Inoltre, presi $\lambda \in \mathbb{R}$ e $v \in V$,

$$\begin{aligned}
(f +_{\mathcal{L}} g)(\lambda v) &= f(\lambda v) +_W g(\lambda v) = \lambda f(v) +_W \lambda g(v) = \\
&= \lambda(f(v) +_W g(v)) = \lambda(f +_{\mathcal{L}} g)(v).
\end{aligned}$$

Quindi $f +_{\mathcal{L}} g$ è un elemento di $\mathcal{L}in(V, W)$ perciò "$+_{\mathcal{L}}$" è un'operazione ben definita su $\mathcal{L}in(V, W)$.

Definiamo ora il prodotto di $\eta \in \mathbb{R}$ per $f \in \mathcal{L}in(V, W)$, $\eta f : V \to W$, ponendo $(\eta f)(v) = \eta(f(v))$. Si vede facilmente che anche ηf è una applicazione lineare e quindi un elemento di $\mathcal{L}in(V, W)$. Si verifichi che $\mathcal{L}in(V, W)$ con le operazioni appena definite è un $\mathbb{R}$-spazio vettoriale.

Scegliamo ora una base $\mathcal{V}$ di V e una base $\mathcal{W}$ di W: $\mathcal{V} = \{v_1, \ldots, v_{n_1}\}$, $\mathcal{W} = \{w_1, \ldots, w_{n_2}\}$. Con queste scelte possiamo definire una funzione

$$\begin{array}{rccc} \varphi_{\mathcal{V},\mathcal{W}}: & \mathcal{L}in(V, W) & \longrightarrow & \mathcal{M}_{n_2,n_1}(\mathbb{R}) \\ & f & \longmapsto & A_f \end{array}$$

che ad ogni applicazione lineare f associa la matrice A_f relativa all'applicazione f rispetto alle basi $\mathcal{V}$ del dominio, $\mathcal{W}$ del codominio. Ricordiamo che tale applicazione dipende dalla scelta delle basi, che le colonne di A_f sono le coordinate rispetto ai vettori della base del codominio delle immagini dei vettori della base del dominio, che ambedue gli spazi $\mathcal{L}in(V, W)$ e $\mathcal{M}_{n_2,n_1}(\mathbb{R})$ ammettono una struttura di spazio vettoriale.

Proposizione 6.1.2 *La funzione $\varphi_{\mathcal{V},\mathcal{W}}$ è un isomorfismo.*

Dimostrazione. Dobbiamo vedere che $\varphi_{\mathcal{V},\mathcal{W}}$ è una funzione lineare biiettiva. Per la linearità basta osservare che la matrice associata a $f +_{\mathcal{L}} g$ ha sulle colonne le coordinate rispetto a $\mathcal{W}$ delle immagini dei vettori della base $\mathcal{V}$ tramite $f +_{\mathcal{L}} g$: nella prima colonna avremmo le coordinate del vettore $(f +_{\mathcal{L}} g)(v_1) = f(v_1) + g(v_1)$ cioè la somma delle coordinate di $f(v_1)$ e di quelle di $g(v_1)$. Ma le coordinate di $f(v_1)$ costituiscono la prima colonna di A_f e le coordinate di $g(v_1)$ la prima colonna di A_g. Dunque la prima colonna della matrice $A_{f+_{\mathcal{L}}g}$, associata a $f +_{\mathcal{L}} g$, è somma della prima colonna di A_f e della prima colonna di A_g. Lo stesso vale per tutti gli altri elementi della base $\mathcal{V}$ e quindi per tutte le altre colonne di $A_{f+_{\mathcal{L}}g}$:

$$\varphi_{\mathcal{V},\mathcal{W}}(f +_{\mathcal{L}} g) = A_{f+_{\mathcal{L}}g} = A_f +_{\mathcal{M}} A_g = \varphi_{\mathcal{V},\mathcal{W}}(f) +_{\mathcal{M}} \varphi_{\mathcal{V},\mathcal{W}}(g).$$

Per quanto riguarda poi $\lambda \in \mathbb{R}$ e $f \in \mathcal{L}in(V, W)$ si ha che λf è per definizione l'applicazione lineare che spedisce ogni vettore v_i in λ-volte $f(v_i)$, i.e. in λ-volte le vecchie coordinate. Otteniamo pertanto $\varphi_{\mathcal{V},\mathcal{W}}(\lambda f) = A_{\lambda f} = \lambda A_f$:

$$\varphi_{\mathcal{V},\mathcal{W}}(\lambda f) = \lambda \varphi_{\mathcal{V},\mathcal{W}}(f).$$

Quindi la funzione $\varphi_{\mathcal{V},\mathcal{W}}$ è lineare. Vediamo che $\varphi_{\mathcal{V},\mathcal{W}}$ è un'applicazione iniettiva: l'elemento di $\mathcal{L}in(V,W)$ che ha come immagine il vettore neutro dello spazio vettoriale delle matrici $\mathcal{M}_{n_2,n_1}(\mathbb{R})$, i.e. la matrice nulla, è l'applicazione lineare h di V in W che manda ogni vettore della base $\mathcal{V}$ nel vettore di W che ha coordinate tutte nulle cioè nel vettore nullo di W. Dunque h è l'applicazione lineare nulla! Per la suriettività osserviamo che, data una matrice $B = (b_{ij}) \in \mathcal{M}_{n_2,n_1}(\mathbb{R})$, possiamo associare a tale matrice la applicazione lineare g_B che spedisce v_1 nel vettore di W che abbia come coordinate rispetto a $\mathcal{W}$ la prima colonna di B: $g_B(v_1) = b_{11}w_1 + b_{21}w_2 + \cdots + b_{n_21}w_{n_2}$ e così via. La matrice associata a g_B rispetto alle basi $\mathcal{V}$ e $\mathcal{W}$ è esattamente B! La funzione $\varphi_{\mathcal{V},\mathcal{W}}$ è dunque suriettiva. **C.V.D.**

Conclusione: scelte una base di V ed una base di W, parlare di applicazioni lineari tra V e W o di matrici ad esse associate (tramite la scelta delle basi) è del tutto equivalente. Per questo andremo a studiare direttamente le matrici. Dall'osservazione precedente possiamo anche dedurre che l'insieme delle applicazioni lineari $\mathcal{L}in(V,W)$ tra due spazi vettoriali di dimensione finita, n e m rispettivamente, è pure uno spazio vettoriale di dimensione mn, perché isomorfo allo spazio delle matrici. Attenzione: tale isomorfismo dipende dalle basi scelte! Non esiste una matrice che rappresenti una applicazione lineare se non dopo aver specificato le basi! Vi sono tuttavia delle caratteristiche intrinseche nella applicazione lineare da cui siamo partiti che si ritroveranno in tutte le matrici ad essa associate rispetto a basi diverse.

6.2 Risolvere i sistemi lineari

Arriviamo al nostro obbiettivo principale: studiare i sistemi lineari (problema A dell'Introduzione). Che cos'è un sistema lineare?

Definizione 6.2.1 *Un sistema lineare di m equazioni in n incognite a coefficienti reali è una lista di m equazioni nelle incognite $x_1, x_2, \ldots, x_n$*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{6.1}$$

dove gli elementi $a_{ij} \in \mathbb{R}$, per $i = 1, \ldots, m$ e $j = 1, \ldots, n$, si dicono i coefficienti del sistema e gli elementi $b_i \in \mathbb{R}$, per $i = 1, \ldots, m$ si dicono i *termini noti del sistema. Il sistema si dice omogeneo se tutti i termini noti sono nulli, e si dice quadrato se $m = n$. La matrice*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

si dice matrice incompleta associata al sistema (6.1) mentre la matrice

$$(A \mid \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

che si ottiene aggiungendo ad A la colonna dei termini noti $\mathbf{b}$ *(quindi la matrice* $(A \mid \mathbf{b}) \in \mathcal{M}_{m,n+1}(\mathbb{R})$*) si dice matrice completa associata a (6.1).*

Per determinare le soluzioni di un sistema lineare sfrutteremo la teoria sviluppata finora.

Abbiamo visto che, dati due spazi vettoriali V e W di dimensione rispettivamente n e m, e scelte due basi $\mathcal{V} = \{v_1, \dots, v_n\}$ e $\mathcal{W} = \{w_1, \dots, w_m\}$ rispettivamente di V e W, un'applicazione lineare $f \in \mathcal{L}in(V, W)$ è rappresentata da una matrice $A = (a_{ij})_{i=1,\dots m}^{j=1,\dots,n} \in \mathcal{M}_{m,n}(\mathbb{R})$ in cui la colonna $j-esima$ è data dalle coordinate, rispetto alla base $\mathcal{W}$, dell'immagine del $j-esimo$ vettore della base $\mathcal{V}$, i.e., $f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m = (a_{1j}, a_{2j}, \dots, a_{mj})_{\mathcal{W}}$. Risulta altresì che per calcolare l'immagine mediante f del vettore $v \in V$ si devono innanzitutto calcolare le sue coordinate nella base $\mathcal{V}$: $v = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n = (\beta_1, \beta_2, \dots, \beta_n)_{\mathcal{V}}$, dopodiché il prodotto $A \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$ fornirà le coordinate nella base $\mathcal{W}$ del vettore $f(v)$ (attenzione: espresse in colonna, non in riga). La scelta delle basi ci è servita solamente per interpretare le applicazioni lineari in termini di matrici.

Prendiamo ora una matrice $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{R})$ ed una $m-upla$ $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$ (colonna dei termini noti). Allora cercare le $n-uple$

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ (colonna delle incognite) tali che $A\mathbf{x} = \mathbf{b}$, cioè

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m-1\,1} & a_{m-1\,2} & a_{m-1\,3} & \dots & \dots & a_{m-1\,n} \\ a_{m1} & a_{m2} & a_{m3} & \dots & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ \dots \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ \dots \\ b_m \end{pmatrix},$$

equivale a cercare i vettori di coordinate $(x_1, x_2, \dots, x_n)$ che sono inviati in $(b_1, b_2, \dots, b_m)$ tramite l'applicazione lineare descritta dalla matrice A. Rispetto a quali basi? Dal momento che stiamo esprimendo tutto in coordinate stiamo fissando implicitamente ed una volta per tutte la base canonica di $\mathbb{R}^n$ nel dominio e la base canonica di $\mathbb{R}^m$ nel codominio. Denotiamo con $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare associata ad A rispetto alle basi canoniche del dominio e del codominio.

A questo punto l'interpretazione del problema iniziale è chiara: determinare tutti i vettori di $\mathbb{R}^n$ la cui immagine tramite l'applicazione lineare φ_A è il vettore $(b_1, \dots, b_m) \in \mathbb{R}^m$. Da 5.1.4 conosciamo già tutti i risultati possibili:

-) se il vettore $(b_1, \dots, b_m) \in \mathbb{R}^m$ non appartiene all'immagine di φ_A il sistema non è risolubile,

-) se $(b_1, \dots, b_m)$ appartiene all'immagine di φ_A, l'insieme delle soluzioni del sistema è $(a_1, a_2, \dots, a_n) + \mathrm{Ker}\varphi_A$ (vedi 5.1) dove $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ è una soluzione particolare del sistema, i.e.,

$$A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Non ci sono altre possibilità.

Teorema 6.2.2 (Teorema di Rouché-Capelli) *Dato un sistema lineare*

$$A\mathbf{x} = \mathbf{b} \quad A \in \mathcal{M}_{m,n}(\mathbb{R}), \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{b} \in \mathbb{R}^m \textit{ (vettori colonna)}$$

esso ammette soluzione se e solo se:

$$\mathrm{rg}A = \mathrm{rg}(A \mid \mathbf{b})$$

cioè se e solo se il rango della matrice incompleta coincide con il rango della matrice completa. Nel caso di uguaglianza l'insieme delle soluzioni è il sottoinsieme $v + \mathrm{Ker}A$ *di* $\mathbb{R}^n$*, ove* v *è una soluzione particolare del sistema e* $\mathrm{Ker}A$ *è il nucleo dell'applicazione lineare* φ_A*. Il sottospazio vettoriale* $\mathrm{Ker}A$ *ha dimensione* $n - \mathrm{rg}A$*.*

Dimostrazione. L'idea è la seguente: nell'interpretazione data del sistema in termini dell'applicazione lineare $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ dire che il sistema ha soluzione significa che il vettore w che ha coordinate $\mathbf{b}$ rispetto alla base canonica di $\mathbb{R}^m$ appartiene all'immagine dell'applicazione φ_A e cioè che la colonna $\mathbf{b}$ è combinazione lineare delle colonne di A (che generano l'immagine di φ_A). Pertanto dire che il sistema ammette soluzione equivale a dire che il numero di colonne linearmente indipendenti della matrice A non cambia se si aggiunge ad A la colonna dei termini noti, cioè che il rango della matrice completa $(A \mid \mathbf{b}) \in \mathcal{M}_{m,n+1}(\mathbb{R})$ è uguale al rango della matrice incompleta A.

La struttura della controimmagine di un vettore tramite una applicazione lineare è già stata studiata nella lezione 5 quindi il resto del contenuto del teorema è già stato dimostrato. **C.V.D.**

Osservazione 6.2.3 Abbiamo già osservato che in uno spazio vettoriale V un sottoinsieme della forma $v + \ker A$ è un sottospazio vettoriale di V se e solo se $v \in \ker A$. In tal caso $v + \ker A = \ker A$. Sia ora $v + \ker A$ l'insieme delle soluzioni di un sistema lineare della forma $A\mathbf{x} = \mathbf{b}$. Allora tale insieme è un sottospazio vettoriale se e solo se il sistema è omogeneo, vale a dire se e solo se $\mathbf{b}$=$\mathbf{0}$.

Osserviamo inoltre che l'insieme $v+\ker A$ è costituito da un solo elemento se $\ker A$ è banale altrimenti è costituito da infiniti elementi, perché se $\ker A$ non è banale contiene infiniti elementi. Dunque se un sistema lineare a coefficienti reali ammette soluzioni, tali soluzioni sono infinite o una sola. Non ci sono altre possibilità.

A questo punto non ci resta che trovare un modo per determinare le soluzioni del sistema qualora questo ammetta soluzioni. Vediamo quello che sappiamo già. Per quanto già osservato, $\mathrm{Ker}A$ è l'insieme delle soluzioni del sistema omogeneo $A\mathbf{x} = \mathbf{0}$ (continuiamo ad indicare i vettori colonna in grassetto). D'altra parte, per 5.2.1, Ker$A$ è un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione $n - \dim(\mathrm{Im}A) = n - \mathrm{rg}A$.

6.3 Il Metodo di soluzione di un sistema lineare

Di fronte ad un sistema lineare $A\mathbf{x} = \mathbf{b}$ ci poniamo le seguenti domande:

1) Il sistema ammette soluzioni?

Se la risposta a questa domanda è negativa allora non abbiamo altro da chiederci. Se la risposta è affermativa ci chiediamo:

2) Quante soluzioni ammette il sistema?
3) Quali sono le soluzioni del sistema?

Abbiamo visto che ad ogni sistema lineare $A\mathbf{x} = \mathbf{b}$ (con $A \in \mathcal{M}_{m,n}$) possiamo associare l'applicazione lineare $\varphi_A : \mathbb{R}^n \to \mathbb{R}^m$ descritta, rispetto alle basi canoniche del dominio e del codominio, dalla matrice A. Abbiamo visto che determinare le soluzioni di un sistema equivale a determinare l'insieme $\varphi_A^{-1}(\{w\})$ ove $w \in \mathbb{R}^m$ è il vettore che ha coordinate $\mathbf{b}$ rispetto alla base canonica di $\mathbb{R}^m$.

Se cambiassimo la base del codominio? Allora la matrice associata a φ_A cambierebbe di conseguenza, ma cambierebbe anche il termine noto $\mathbf{b}$ del sistema che risulterebbe uguale alla m-upla di coordinate del vettore w rispetto alla nuova base di $\mathbb{R}^m$ scelta. In termini di sistemi lineari questo significa risolvere un sistema diverso da quello dato, ma *equivalente* cioè con le stesse soluzioni del precedente, infatti il vettore w è sempre lo stesso e l'insieme delle soluzioni è $\varphi^{-1}(\{w\})$.

In particolare, il rango della matrice non sarà cambiato! Infatti modificheremo solo la base del codominio e dunque la dimensione dell'immagine non verrà alterata. Ma quali effetti produce un cambiamento di base del codominio sulla matrice A?

Mostreremo in 8.4 che scambiare due vettori della base del codominio equivale a scambiare le righe corrispondenti della matrice A. Moltiplicare il vettore $i-esimo$ della base del codominio per $\alpha \neq 0$ equivale a moltiplicare la riga $i-esima$ per $\frac{1}{\alpha}$. Scegliere una nuova base in cui, ad esempio, l'$i-esimo$ vettore è il vecchio $i-esimo$ a cui si somma β volte il $j-esimo$ e gli altri rimangono invariati corrisponde a sommare alla riga $j-esima$ della matrice A la riga $i-esima$ moltiplicata per $-\beta$. Come si vede, i cambiamenti di base nel codominio determinano trasformazioni che coinvolgono le righe della matrice A. Tali trasformazioni non alterano il rango della matrice in quanto l'immagine dell'applicazione lineare φ_A è sempre la stessa. Tali operazioni sono dette operazioni elementari sulle righe:

Definizione 6.3.1 Operazioni elementari sulle righe. *Data una matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$, le seguenti operazioni si dicono "operazioni elementari sulle righe di" A:*

(i) *moltiplicare una riga per uno scalare non nullo;*

(ii) *sommare ad una riga un'altra riga moltiplicata per uno scalare;*

(iii) *scambiare di posto tra di loro due righe.*

Notazione:

indicheremo con $H_i(\alpha)$ l'operazione elementare (i): moltiplicare la riga i-esima per lo scalare non nullo $\alpha \neq 0$;

indicheremo con $H_{i,j}(\alpha)$ l'operazione elementare (ii): sommare alla riga i-esima la riga j-esima moltiplicata per α;

indicheremo con $S_{i,j}$ l'operazione elementare (iii): scambiare la riga i-esima con la riga j-esima.

Osservazione 6.3.2 Le operazioni elementari sulle righe di una matrice non alterano il rango della matrice.

L'idea è quella di usare i cambiamenti di base nel codominio e quindi operazioni elementari sulle righe in modo da ottenere un sistema lineare equivalente di cui sia facile determinare l'insieme delle soluzioni.

Definizione 6.3.3 *Una matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ si dice a* **scala (o a scalini) rispetto alle righe** *se soddisfa le seguenti condizioni: le righe che contengono solo zeri sono le ultime e il primo elemento non nullo di ogni riga (non nulla) si trova più a destra del primo elemento non nullo della riga precedente.*

Ad esempio

$$\begin{pmatrix} \mathbf{1} & 0 & 3 & 0 & 0 & 0 \\ 0 & \mathbf{1} & 2 & 4 & 1 & 0 \\ 0 & 0 & 0 & \mathbf{2} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

è una matrice a scala rispetto alle righe.

Il primo elemento non nullo di una riga di una matrice a scala si chiama anche pivot.

Perché ci interessa questo tipo di matrice? Perché in una matrice a scalini rispetto alle righe i pivots individuano esattamente le colonne linearmente indipendenti, quindi il rango colonne di una matrice in forma a scalini rispetto alle righe è uguale al numero di pivots (per convincersene, basta osservare che le colonne che non contengono pivots sono combinazioni lineari delle colonne alla loro sinistra che contengono pivots!) Nel nostro esempio abbiamo, per la quinta colonna,

$$\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ 4 \\ 2 \\ 0 \\ 0 \end{pmatrix} + (-1) \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Tra le righe...osserviamo che possiamo definire il *rango righe* della matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ come il massimo numero di righe di A (intese come vettori di $\mathbb{R}^n$) linearmente indipendenti. Quello che si può vedere è che i pivots nella forma a scalini rispetto alle righe appartengono esattamente alle righe che sono linearmente indipendenti. La conclusione è abbastanza sorprendente: data una matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$, le colonne di A generano un sottospazio di $\mathbb{R}^m$ e le righe generano un sottospazio di $\mathbb{R}^n$: tali sottospazi hanno la stessa dimensione! D'ora in poi parleremo semplicemente di *rango di una matrice* A, senza specificare se si tratta del rango righe o del rango colonne, e lo indicheremo con $\mathrm{rg}A$.

Le operazioni elementari sulle righe di una matrice lasciano invariato l'insieme delle soluzioni del nostro sistema lineare, dal momento che la base del dominio non viene coinvolta. Stiamo prendendo due piccioni con una fava: semplifichiamo il nostro sistema e otteniamo immediatamente quanto vale $\mathrm{rg}A$ e quindi quanto vale la dimensione del nucleo della matrice A ($\dim \mathrm{Ker}A = n - \mathrm{rg}A$).

Ma è sempre possibile "ridurre" una matrice $A \in \mathcal{M}_{m,n}$ in forma a scalini e, se sì, come? Mostreremo che con operazioni elementari sulle righe si può sempre "ridurre" una matrice in forma a scalini utilizzando il Teorema di riduzione di Gauss.

Teorema 6.3.4 (**Teorema di riduzione di Gauss**). *Ogni matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$ può essere ridotta in forma a scalini (per righe) usando operazioni elementari sulle righe di A.*

Illustriamo concretamente il **metodo di riduzione di Gauss** in un esempio.

Riduciamo a scala la matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 2 & 0 \\ 2 & 0 & 0 & -1 \end{pmatrix}.$$

Passo 1. Individuare la prima colonna non nulla. Nel nostro caso è la prima colonna. Nella prima colonna non nulla individuare un'entrata non nulla. Se tale entrata non è nella prima riga, con un'operazione elementare del tipo (iii) (scambio di due righe), spostare tale riga al posto della prima. Nel nostro esempio $a_{11} = a_{21} = 0$, quindi la prima entrata non nulla della prima colonna è $a_{31} = 1$. Allora scambiamo la prima con la terza riga:

$$A = \begin{pmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 2 & 0 \\ 2 & 0 & 0 & -1 \end{pmatrix} \xrightarrow{S_{1,3}} \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 0 & -1 \end{pmatrix}.$$

Passo 2. Ora tramite operazioni elementari del secondo tipo (sommando ad ogni riga un opportuno multiplo della nostra nuova prima riga) possiamo fare in modo che tutti gli altri termini di quella colonna diventino nulli. Nel nostro esempio $a_{41} = 2$, quindi (essendo ora $a_{11} = 1$) per avere 0 nel posto $4,1$ dobbiamo sommare alla quarta riga la prima moltiplicata per -2:

$$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 0 & -1 \end{pmatrix} \xrightarrow{H_{4,1}(-2)} \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 2 & -4 & -1 \end{pmatrix}.$$

Passo 3. Lasciare fissa la prima riga e procedere come ai passi 1 e 2 per la matrice che si ottiene cancellando la prima riga.

Nel nostro esempio considerare la matrice che si ottiene cancellando la prima riga vuol dire che il primo termine non nullo della seconda colonna è $a_{23} = 1$ (perché la prima riga non va considerata e $a_{22} = 0$). Allora, procedendo come spiegato nel passo 1, scambiamo la seconda riga con la terza:

$$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 2 & -4 & -1 \end{pmatrix} \xrightarrow{S_{2,3}} \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 2 & -4 & -1 \end{pmatrix}.$$

Nella seconda colonna (sotto la seconda riga) resta ancora un termine non nullo $a_{42} = 2$. Per eliminarlo dobbiamo sommare alla quarta riga la seconda moltiplicata per -2:

$$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 2 & -4 & -1 \end{pmatrix} \xrightarrow{H_{4,2}(-2)} \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -4 & -7 \end{pmatrix}.$$

Ora passiamo alla terza colonna lasciando fisse le prime due righe. Il coefficiente $a_{33} = -1$ quindi in questo caso non serve fare nessuno scambio. Ci resta da "sistemare" solo l'ultima riga, in quanto $a_{43} = -4$. In questo caso basterà sommare all'ultima riga la terza moltiplicata per -4:

$$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -4 & -7 \end{pmatrix} \xrightarrow{H_{4,3}(-4)} \begin{pmatrix} \mathbf{1} & -1 & 2 & 0 \\ 0 & \mathbf{1} & 0 & 3 \\ 0 & 0 & \mathbf{-1} & 1 \\ 0 & 0 & 0 & \mathbf{-11} \end{pmatrix}.$$

In questo modo abbiamo ridotto la nostra matrice di partenza A in forma a scalini e i termini in neretto nella matrice ottenuta sono i pivots.

A questo punto abbiamo tutti gli strumenti per rispondere alle domande che ci siamo posti all'inizio della sezione 6.3:

Risoluzione di un sistema lineare

Sia dato il sistema lineare $A\mathbf{x} = \mathbf{b}$ con $A \in \mathcal{M}_{m,n}(\mathbb{R})$ e $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$:

1) Applichiamo le trasformazioni per riga alla matrice completa $(A \mid \mathbf{b})$ fino a ridurla in forma a scalini per righe. Chiamiamo S' la matrice ottenuta.

In questo modo otteniamo contemporaneamente anche la forma a scalini per righe di A che indicheremo con S. Il sistema ha soluzione se e solo se il rango di S', cioè il numero di righe non nulle o, equivalentemente, di pivots di S', è uguale a quello di S, cioè se e solo se S' non ha pivots nell'ultima colonna. Se invece il numero di righe non nulle di S' è diverso da quello di S allora necessariamente S' ha un pivots in più (nell'ultima colonna) rispetto alla matrice S e in questo caso il sistema lineare non ha soluzione.

2) Nel caso in cui vi siano soluzioni possiamo cercarle direttamente usando le matrici S e S': infatti le trasformazioni per riga non hanno alterato le soluzioni. Dobbiamo trovare una soluzione particolare del sistema e determinare $\mathrm{Ker}A$ che è uno spazio di dimensione $n - \mathrm{rg}A$: quindi dovremo trovare $n - \mathrm{rg}A$ vettori linearmente indipendenti in $\mathrm{Ker}A$. Le coordinate dei vettori di $\mathrm{Ker}A$ debbono verificare il sistema omogeneo $A\mathbf{x} = \mathbf{0}$. Tale sistema è equivalente al sistema $S\mathbf{x} = \mathbf{0}$ che si può risolvere per sostituzioni successive dal basso, dal momento che la matrice S è in forma a scala per righe. Analogamente, per trovare una soluzione particolare del sistema di partenza possiamo usare il sistema associato alla matrice S' che è equivalente a quello di partenza. Descriviamo un metodo meccanico per costruire le soluzioni: consideriamo una $n - upla$ $(x_1, \ldots, x_n)$ e lasciamo libere le entrate la cui posizione corrisponde alle colonne dove compaiono i pivots, ad esempio se i pivots compaiono nella prima, quarta, quinta colonna, allora lasciamo libere le entrate x_1, x_4, x_5. Alle altre $n - \mathrm{rg}A$ entrate diamo di volta in volta i seguenti valori: le fissiamo tutte uguali a 0 tranne una che poniamo uguale a 1. Per ognuna di queste scelte (sono $n - \mathrm{rg}A$) possiamo determinare in modo univoco i valori delle variabili che abbiamo lasciato libere. In questo modo abbiamo costruito $n - \mathrm{rg}A$ vettori di $\mathbb{R}^n$ linearmente indipendenti (ognuno di essi ha una entrata non nulla in una posizione in cui tutti gli altri hanno entrate nulle). Ognuna di queste n-uple è soluzione del sistema omogeneo $S\mathbf{x} = \mathbf{0}$ pertanto abbiamo determinato una base dello spazio delle soluzioni del sistema omogeneo. Resta da trovare una soluzione particolare di $A\mathbf{x} = \mathbf{b}$ e per questo è conveniente utilizzare la forma a scalini per righe già ottenuta. In questo caso non esiste un modo standard di procedere. Un'idea può essere quella di assegnare il valore 0 alle incognite x_k relative alle colonne che non contengono pivots e di calcolare poi la soluzione per sostituzione diretta.

Attenzione. Non esiste una forma a scalini per righe unica per una matrice. Tuttavia ogni forma a scalini per righe della stessa matrice possiede lo stesso rango!

Esempio 6.3.5 Risolvere il seguente sistema nelle incognite x_1, x_2, x_3:

$$\begin{cases} 2x_1 - x_2 = 1 \\ x_2 = -3 \end{cases}.$$

Attenzione: nelle equazioni del sistema non appare l'incognita x_3 ma le soluzioni debbono essere cercate tra gli elementi di $\mathbb{R}^3$. In termini matriciali si tratta di risolvere il sistema

$$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix}.$$

Una soluzione particolare è data da $(-1, -3, 1)$ (ma anche $(-1, -3, 50)$ va bene!). Determiniamo ora Ker$A$. Si vede subito che il rango di A è due e che quindi la dimensione di Ker$A$ è uno. La matrice è già in forma a scalini per righe, dunque scegliamo il valore $x_3 = 1$ e cerchiamo un elemento di Ker$A$ della forma $(x_1, x_2, 1)$. Otteniamo facilmente il vettore $(0, 0, 1)$. Quindi Ker$A = \langle(0, 0, 1)\rangle$ e l'insieme delle soluzioni del sistema è $(1, -3, 0) + \langle(0, 0, 1)\rangle$.

Esempio 6.3.6 Cerchiamo di risolvere il sistema lineare nelle incognite x_1, x_2, x_3, x_4:

$$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1 \\ 5x_2 - 3x_3 - 5x_4 = -3 \\ 3x_1 - 4x_2 + 3x_3 + 4x_4 = 3. \end{cases}$$

Tale sistema può essere scritto in forma matriciale come

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \\ 3 & -4 & 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix},$$

i.e. $A\mathbf{x} = \mathbf{b}$ dove $A = \begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \\ 3 & -4 & 3 & 4 \end{pmatrix} \in \mathcal{M}_{3,4}(\mathbb{R})$ si deve intendere come la matrice di una applicazione lineare da $\mathbb{R}^4$ a $\mathbb{R}^3$. In particolare A avrà necessariamente nucleo non banale poiché la dimensione dell'immagine è al più 3! Prendiamo allora la matrice completa del sistema e cerchiamo di trovarne la forma a scalini per righe. Sostituiamo l'ultima riga con la somma della prima riga moltiplicata per $-\frac{3}{2}$ e dell'ultima (operazione elementare $H_{3,1}(-3/2)$). Si ottiene:

$$\left(\begin{array}{cccc|c} 2 & -1 & 1 & 1 & 1 \\ 0 & 5 & -3 & -5 & -3 \\ 0 & -\frac{5}{2} & \frac{3}{2} & \frac{5}{2} & \frac{3}{2} \end{array}\right).$$

Ora sostituiamo la terza riga con la somma della seconda riga moltiplicata per $\frac{1}{2}$ e dell'ultima:

$$\left(\begin{array}{cccc|c} 2 & -1 & 1 & 1 & 1 \\ 0 & 5 & -3 & -5 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

che è la forma a scalini per riga cercata. In particolare abbiamo calcolato la forma a scalini della matrice A:

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il numero di pivots è due quindi la dimensione dell'immagine dell'applicazione è 2, la dimensione del nucleo è $n - \mathrm{rg}A = 4-2 = 2$. Il rango della matrice completa è pure 2 (non ci sono pivots nell'ultima colonna): il sistema ammette dunque soluzioni. L'insieme delle soluzioni è $(a_1, a_2, a_3, a_4) + \mathrm{Ker}A$, essendo (a_1, a_2, a_3, a_4) una soluzione particolare del sistema. Determiniamo $\mathrm{Ker}A$: i pivots sono nella prima e seconda colonna. La terza e la quarta colonna della matrice A non contengono pivots. Cerchiamo allora gli elementi di $\mathrm{Ker}A$ del tipo $(x_1, x_2, 1, 0)$ e $(x_1, x_2, 0, 1)$. Risolviamo quindi, innanzitutto, il sistema:

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{cases} 2x_1 - x_2 + 1 + 0 = 0 \\ 5x_2 - 3 + 0 = 0 \end{cases}.$$

Otteniamo: $x_2 = \frac{3}{5}$ e $x_1 = \frac{-1}{5}$. Abbiamo così individuato il vettore $(-\frac{1}{5}, \frac{3}{5}, 1, 0)$ appartenente a $\mathrm{Ker}A$. Analogamente possiamo procedere con

$(x_1, x_2, 0, 1)$ e quindi risolvere il sistema

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{cases} 2x_1 - x_2 + 0 + 1 = 0 \\ 5x_2 + 0 - 5 = 0 \end{cases}.$$

Otteniamo $x_2 = 1$ e $x_1 = 0$, cosicché la seconda soluzione è data da $(0, 1, 0, 1)$. Quindi $\mathrm{Ker}A = \langle(0, 1, 0, 1), (-\frac{1}{5}, \frac{3}{5}, 1, 0)\rangle$ cioè $\mathrm{Ker}A$ è l'insieme dei vettori di $\mathbb{R}^4$ della forma $(-\frac{\lambda}{5}, \frac{3\lambda}{5} + \eta, \lambda, \eta)$ al variare di η e λ nell'insieme dei numeri reali. Ora resta da trovare una soluzione particolare del sistema completo usando la forma a scalini per righe che abbiamo determinato:

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & 5 & -3 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

cioè

$$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 1 \\ 5x_2 - 3x_3 - 5x_4 = -3 \end{cases}.$$

Se poniamo $x_3 = x_4 = 0$ otteniamo la soluzione $x_2 = -\frac{3}{5}$ e $x_1 = \frac{1}{5}$. L'insieme completo delle soluzioni del sistema è dunque dato da

$$\Big(\frac{1}{5}, -\frac{3}{5}, 0, 0\Big) + \langle(0, 1, 0, 1), \Big(-\frac{1}{5}, \frac{3}{5}, 1, 0\Big)\rangle =$$
$$= \Big\{\Big(\frac{1}{5} - \frac{\lambda}{5}, -\frac{3}{5} + \frac{3\lambda}{5} + \eta, \lambda, \eta\Big) \mid \eta, \lambda \in \mathbb{R}\Big\}.$$

Esempio 6.3.7 *Risoluzione di un sistema lineare parametrico.* Consideriamo, al variare di $\alpha \in \mathbb{R}$, il sistema nelle incognite x, y, z:

$$\mathcal{S}_\alpha = \begin{cases} \alpha x + y + 2z = 1 \\ x - 2y - \alpha z = 2 \\ 3x - y = 3 \end{cases}$$

e studiamone le soluzioni per ogni α. Alla fine dell'esercizio dovremo essere in grado di rispondere, ad esempio, alla domanda: "Se $\alpha = 1$ qual è l'insieme delle soluzioni del sistema? E se $\alpha = 5$? E se $\alpha = \sqrt{2}$?" Possiamo pensare al parametro α come ad una variabile temporale. Allora risolvere il sistema omogeneo associato a S_α, al variare di α, significa studiare il nucleo di un operatore che si evolve nel tempo. L'idea è quella di studiare come varia il nucleo dell'operatore nel tempo per poi essere in grado di dire che cosa succede in un qualsiasi istante fissato.

Procediamo con la soluzione. Il problema consiste nel determinare la controimmagine del vettore $(1, 2, 3)$ mediante le applicazioni lineari di matrice

$$A_\alpha = \begin{pmatrix} \alpha & 1 & 2 \\ 1 & -2 & -\alpha \\ 3 & -1 & 0 \end{pmatrix}.$$

A seconda del valore di α il vettore $(1, 2, 3)$ starà nell'immagine dell'applicazione o meno. Consideriamo la matrice completa del sistema e, ricordando che α è un numero reale, effettuiamo delle trasformazioni per riga che, come sappiamo, non alterano le soluzioni del sistema:

$$\left(\begin{array}{ccc|c} \alpha & 1 & 2 & 1 \\ 1 & -2 & -\alpha & 2 \\ 3 & -1 & 0 & 3 \end{array}\right) \xrightarrow{S_{1,2}} \left(\begin{array}{ccc|c} 1 & -2 & -\alpha & 2 \\ \alpha & 1 & 2 & 1 \\ 3 & -1 & 0 & 3 \end{array}\right) \xrightarrow{H_{3,1}(-3)} \left(\begin{array}{ccc|c} 1 & -2 & -\alpha & 2 \\ \alpha & 1 & 2 & 1 \\ 0 & 5 & 3\alpha & -3 \end{array}\right) \rightarrow$$

$$\xrightarrow{H_{2,1}(-\alpha)} \left(\begin{array}{ccc|c} 1 & -2 & -\alpha & 2 \\ 0 & 2\alpha+1 & 2+\alpha^2 & 1-2\alpha \\ 0 & 5 & 3\alpha & -3 \end{array}\right) \xrightarrow{S_{3,2}} \left(\begin{array}{ccc|c} 1 & -2 & -\alpha & 2 \\ 0 & 5 & 3\alpha & -3 \\ 0 & 2\alpha+1 & 2+\alpha^2 & 1-2\alpha \end{array}\right) \rightarrow$$

$$\xrightarrow{H_3(5)H_{3,2}(-(2\alpha+1)/5)} \left(\begin{array}{ccc|c} 1 & -2 & -\alpha & 2 \\ 0 & 5 & 3\alpha & -3 \\ 0 & 0 & -\alpha^2-3\alpha+10 & -4\alpha+8 \end{array}\right).$$

Il rango della matrice incompleta che abbiamo ottenuto è uguale a 3 per tutti gli α tali che $-\alpha^2 - 3\alpha + 10 \neq 0$, cioè per ogni $\alpha \neq 2, -5$. Pertanto per $\alpha \neq 2, -5$ la matrice incompleta del sistema ha rango 3 e dunque risulta la matrice di un isomorfismo (essendo la matrice di un endomorfismo ed essendo suriettiva) e quindi il sistema ammette una ed una sola soluzione. Il sistema di partenza è equivalente al sistema:

$$\begin{cases} x - 2y - \alpha z = 2 \\ 5y + 3\alpha z = -3 \\ (-\alpha^2 - 3\alpha + 10)z = -4\alpha + 8 \end{cases}$$

che ha come unica soluzione il vettore $(\frac{4\alpha-8}{\alpha^2+3\alpha-10}, \frac{-3\alpha^2+3\alpha+6}{\alpha^2+3\alpha-10}, \frac{4\alpha-8}{\alpha^2+3\alpha-10})$.

Se $\alpha = 2$ oppure $\alpha = -5$ la matrice incompleta ha rango 2. Qual è, in questi casi, il rango della matrice completa? Per $\alpha = 2$ la matrice completa è:

$$\left(\begin{array}{ccc|c} 1 & -2 & -2 & 2 \\ 0 & 5 & 6 & -3 \\ 0 & 0 & 0 & -4(2)+8 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -2 & -2 & 2 \\ 0 & 5 & 6 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

pertanto sia la matrice completa che la matrice incompleta hanno rango due. Il sistema ammette soluzioni; l'insieme delle soluzioni è della forma $(a, b, c) + KerA_2$ e il nucleo di A_2 è un sottospazio di $\mathbb{R}^3$ di dimensione 1. Il sistema di partenza è equivalente al sistema

$$\begin{cases} x - 2y - 2z = 2 \\ 5y + 6z = -3 \end{cases}$$

e l'insieme delle soluzioni è $(1, 0, -\frac{1}{2}) + \langle(2, 6, -5)\rangle$. Se $\alpha = -5$ la forma a scalini per righe è:

$$\left(\begin{array}{ccc|c} 1 & -2 & 5 & 2 \\ 0 & 5 & -15 & -3 \\ 0 & 0 & 0 & -4(-5)+8 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & -2 & 5 & 2 \\ 0 & 5 & -15 & -3 \\ 0 & 0 & 0 & 28 \end{array}\right).$$

In questo caso il sistema non ammette soluzioni poiché il rango della matrice incompleta è diverso dal rango della matrice completa.

Osservazione importante. Supponiamo di dover studiare, al variare del parametro reale a, il sistema

$$\begin{cases} ax + y = 1 \\ x + z = 0. \end{cases}$$

La matrice completa associata al sistema dato è la matrice

$$A = \left(\begin{array}{ccc|c} a & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right).$$

Si tratta di ridurre la matrice A in forma a scalini per righe.

COME PROCEDERE: scambiamo le righe di A in modo che l'elemento di posto 1,1 NON contenga il parametro a:

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ a & 1 & 0 & 1 \end{array}\right);$$

sostituiamo alla seconda riga della matrice ottenuta la somma della seconda riga per la prima moltiplicata per $-a$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -a & 1 \end{array}\right).$$

La matrice incompleta e la matrice completa del sistema hanno, al variare di a, lo stesso rango (uguale a 2), pertanto il sistema di partenza ha infinite soluzioni della forma: $(-t, at+1, t)$, con $t \in \mathbb{R}$.

COME NON PROCEDERE: sostituiamo alla seconda riga di A la somma della prima riga per la seconda moltiplicata per $-a$:

$$\left(\begin{array}{ccc|c} a & 1 & 0 & 1 \\ 0 & 1 & -a & 1 \end{array}\right).$$

Anche in questo caso la matrice incompleta e la matrice completa hanno lo stesso rango ma tale rango è uguale ad 1 se $a = 0$. In questo caso il sistema sembrerebbe avere rango 1 ma questo è falso. Se, infatti, sostituiamo $a = 0$ nel sistema dato, otteniamo le due equazioni indipendenti $y = 1$ e $x + z = 0$, cioè un sistema di rango 2. Qual è l'errore nel nostro secondo modo di procedere? Abbiamo sbagliato a sostituire la seconda riga di A con la somma della prima riga per la seconda moltiplicata per $-a$. Infatti nel caso in cui $a = 0$ questo significa cancellare la seconda equazione del sistema (moltiplicandola per zero)!!

6.4 Esercizi svolti

Esercizio 6.4.1 Determinare una base dello spazio vettoriale delle applicazioni lineari di $\mathbb{R}^2$ in $\mathbb{R}^3$: $\mathcal{L}in(\mathbb{R}^2, \mathbb{R}^3)$.

Svolgimento. Sappiamo da 5.3 che, una volta fissate una base di $\mathbb{R}^2$ ed una di $\mathbb{R}^3$, lo spazio vettoriale reale $\mathcal{L}in(\mathbb{R}^2, \mathbb{R}^3)$ è isomorfo allo spazio vettoriale reale delle matrici 3×2 a coefficienti in $\mathbb{R}$ che pertanto ha dimensione 6. Dobbiamo dunque costruire 6 applicazioni lineari da $\mathbb{R}^2$ in $\mathbb{R}^3$ linearmente indipendenti. Fissiamo $\mathcal{E} = \{e_1 = (1,0), e_2 = (0,1)\}$ base di $\mathbb{R}^2$ e $\mathcal{E}' = \{E_1 = (1,0,0), E_2 = (0,1,0), E_3 = (0,0,1)\}$ base di $\mathbb{R}^3$. Consideriamo le applicazioni lineari $\varphi_{ij} \in \mathcal{L}in(\mathbb{R}^2, \mathbb{R}^3)$, con $i = 1,2,3$, $j = 1,2$, così definite: $\varphi_{ij}(e_k) = \delta_{jk} E_i$. Ad esempio, la funzione φ_{12} è la applicazione lineare che manda e_1 nel vettore $\mathbf{0}_{\mathbb{R}^3}$ ed e_2 nel vettore E_1. Mostriamo che, al variare di i, j, le applicazioni φ_{ij} sono linearmente indipendenti: sia dunque

$$\sum_{i,j} a_{ij}\varphi_{ij} = \mathbf{0}_{\mathcal{L}in(\mathbb{R}^2,\mathbb{R}^3)}.$$

Questo significa che per ogni vettore $v \in \mathbb{R}^2$

$$\sum_{i,j} a_{ij}\varphi_{ij}(v) = \mathbf{0}_{\mathbb{R}^3}.$$

In particolare $\sum_{i,j} a_{ij}\varphi_{ij}(e_1) = \mathbf{0}_{\mathbb{R}^3}$, vale a dire, usando la definizione delle mappe φ_{ij},

$$\sum_{ij} a_{ij}\delta_{j1}E_i = \sum_i a_{i1}E_i = 0.$$

Dal momento che i vettori E_i sono linearmente indipendenti, otteniamo allora $a_{i1} = 0$ per ogni i.

Analogamente, $\mathbf{0}_{\mathbb{R}^3} = \sum_{i,j} a_{ij}\varphi_{ij}(e_2) = \sum_{i,j} a_{ij}\delta_{j2}E_i = \sum_i a_{i2}E_i$, da cui deduciamo $a_{i2} = 0$ per ogni i.

Abbiamo così mostrato che le applicazioni φ_{ij} con $i = 1, 2, 3$ e $j = 1, 2$ sono linearmente indipendenti e quindi costituiscono una base di $\mathcal{L}in(\mathbb{R}^2, \mathbb{R}^3)$ (perché sappiamo già che $\dim(\mathcal{L}in(\mathbb{R}^2, \mathbb{R}^3)) = 6$).

Osserviamo che quanto provato equivale a dimostrare che le matrici

$$F_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad F_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad F_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$F_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad F_{31} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad F_{32} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$$

sono una base dello spazio vettoriale $\mathcal{M}_{3,2}(\mathbb{R})$. Notiamo, inoltre, che per ogni $i = 1, 2, 3$ e per ogni $j = 1, 2$, F_{ij} è proprio la matrice associata all'applicazione lineare φ_{ij} rispetto alla base canonica di $\mathbb{R}^2$ e alla base canonica di $\mathbb{R}^3$.

Esercizio 6.4.2 Risolvere il sistema lineare:

$$\begin{cases} 2x - 5y + 12z = 0 \\ 4x - z = 0 \\ 6x - 5y = -1 \end{cases}$$

Svolgimento. Il sistema lineare è un sistema lineare non omogeneo nelle variabili x, y, z. Se esso ammette soluzioni l'insieme delle soluzioni sarà della forma $(a, b, c) + W$ essendo (a, b, c) una soluzione particolare del sistema e W il sottospazio vettoriale di $\mathbb{R}^3$ delle soluzioni del sistema lineare omogeneo

$$\begin{cases} 2x - 5y + 12z = 0 \\ 4x - z = 0 \\ 6x - 5y = 0. \end{cases}$$

La matrice completa associata al sistema è la matrice

$$A = \left(\begin{array}{ccc|c} 2 & -5 & 12 & 0 \\ 4 & 0 & -1 & 0 \\ 6 & -5 & 0 & -1 \end{array} \right).$$

Effettuiamo delle operazioni per riga in modo da ridurre la matrice A in forma a scalini per righe:

$$\left(\begin{array}{ccc|c} 2 & -5 & 12 & 0 \\ 4 & 0 & -1 & 0 \\ 6 & -5 & 0 & -1 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|c} 2 & -5 & 12 & 0 \\ 0 & 10 & -25 & 0 \\ 0 & 10 & -36 & -1 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|c} 2 & -5 & 12 & 0 \\ 0 & 10 & -25 & 0 \\ 0 & 0 & 11 & 1 \end{array} \right)$$

$$\Rightarrow \left(\begin{array}{ccc|c} 2 & -5 & 12 & 0 \\ 0 & 2 & -5 & 0 \\ 0 & 0 & 11 & 1 \end{array} \right).$$

Il rango della matrice completa è uguale a 3 e così pure il rango della matrice incompleta. Di conseguenza la matrice incompleta ha nucleo banale e il sistema ha un'unica soluzione che siamo in grado di calcolare immediatamente usando la forma a scalini per righe che abbiamo determinato: la soluzione del sistema è il vettore $(\frac{1}{44}, \frac{5}{22}, \frac{1}{11})$.

Esercizio 6.4.3 Studiare, al variare del parametro reale h, le soluzioni del sistema nelle incognite x, y, z:

$$\begin{cases} x + hy + z = 1 \\ 2hx + 2z = 0 \\ (1 - h)x + hy = h. \end{cases}$$

Svolgimento. Consideriamo la matrice completa associata al sistema e riduciamola in forma a scalini per righe:

$$\left(\begin{array}{ccc|c} 1 & h & 1 & 1 \\ 2h & 0 & 2 & 0 \\ 1-h & h & 0 & h \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & h & 1 & 1 \\ 0 & -2h^2 & 2-2h & -2h \\ h & 0 & 1 & 1-h \end{array}\right) \Rightarrow$$

$$\Rightarrow \left(\begin{array}{ccc|c} 1 & h & 1 & 1 \\ 0 & -2h^2 & 2-2h & -2h \\ 0 & -h^2 & 1-h & 1-2h \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & h & 1 & 1 \\ 0 & -2h^2 & 2-2h & -2h \\ 0 & 0 & 0 & -2+2h \end{array}\right).$$

Confrontiamo il rango della matrice incompleta con il rango della matrice completa: se $h \neq 1$ il rango della matrice completa è sicuramente diverso dal rango della matrice incompleta, pertanto il sistema non ammette soluzioni.

Se $h = 1$ la matrice completa che abbiamo ottenuto è:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

In questo caso la matrice completa e la matrice incompleta hanno entrambe rango 2: il sistema ammette pertanto soluzioni. Il sistema di partenza è equivalente al sistema

$$\begin{cases} x+y+z=1 \\ -2y=-2 \end{cases}$$

a cui è associato il sistema omogeneo:

$$\begin{cases} x+y+z=0 \\ -2y=0 \end{cases}.$$

Lo spazio delle soluzioni del sistema omogeneo è dunque $\langle(1,0,-1)\rangle$ e $(1,1,-1)$ è una soluzione particolare del sistema di partenza. L'insieme delle soluzioni è dunque: $(1,1,-1)+\langle(1,0,-1)\rangle$.

Esercizio 6.4.4 Si scriva un sistema lineare che abbia come insieme di soluzioni il sottospazio vettoriale S di $\mathbb{R}^3$ generato dai vettori $(3,2,1)$ e $(2,1,1)$. Il sistema richiesto è unico? In caso di risposta negativa se ne scriva un altro.

Svolgimento. Poiché l'insieme delle soluzioni del sistema che stiamo cercando è un sottospazio vettoriale di $\mathbb{R}^3$, il sistema che dobbiamo costruire è un sistema omogeneo del tipo

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

nelle variabili che indicheremo con x_1, x_2 e x_3 e con $A \in \mathcal{M}_{n\times 3}(\mathbb{R})$: il numero delle equazioni che stiamo cercando è al momento incognito e indicato con n. Notiamo che, essendo S un sottospazio di dimensione 2 ed essendo il numero delle variabili uguale a tre, il rango della matrice A deve essere uguale a 1. Il sistema che cerchiamo dovrà perciò essere costituito da una sola equazione indipendente: $ax_1 + bx_2 + cx_3 = 0$. Affinché i vettori $(3, 2, 1)$ e $(2, 1, 1)$ siano soluzioni di questa equazione deve essere:

$$\begin{cases} 3a + 2b + c = 0 \\ 2a + b + c = 0 \end{cases}$$

cioè

$$\begin{cases} a + b = 0 \\ a + c = 0 \end{cases}.$$

Il sistema trovato ha come soluzioni tutti e soli i vettori $(a, b, c) = \lambda(1, -1, -1)$ con $\lambda \in \mathbb{R} \setminus \{0\}$. Dunque, posto $a = 1$, $b = c = -1$, otteniamo l'equazione

$$x_1 - x_2 - x_3 = 0 \tag{6.2}$$

il cui insieme di soluzioni è esattamente $S = \langle (3, 2, 1), (2, 1, 1)\rangle$.

Naturalmente aggiungendo all'equazione (6.2) equazioni ad essa equivalenti si ottengono sistemi equivalenti:

$$\begin{cases} x_1 - x_2 - x_3 = 0 \\ 3x_1 - 3x_2 - 3x_3 = 0. \end{cases}$$

Esercizio 6.4.5 (i) Si scriva, se possibile, un sistema lineare di due equazioni le cui soluzioni siano tutti e soli gli elementi dell'insieme $S = (-1, 1, -3, 0) + \langle (1, 0, 0, -1), (0, 1, -5, 2)\rangle$.

(ii) Si scriva, se possibile, un sistema lineare le cui soluzioni siano tutti e soli gli elementi dell'insieme $T = (2, 0, -1) + \langle (3, 5, 10), (18, 10, 14)\rangle$.

Svolgimento.

(i) Stabiliamo innanzitutto se S è un sottospazio vettoriale o meno, controlliamo cioè se il vettore $(-1, 1, -3, 0)$ appartiene al sottospazio vettoriale generato dai vettori $(1, 0, 0, -1), (0, 1, -5, 2)$. Si tratta di stabilire se esistono $a, b \in \mathbb{R}$ tali che sia

$$(-1, 1, -3, 0) = a(1, 0, 0, -1) + b(0, 1, -5, 2) = (a, b, -5b, -a + 2b).$$

Otteniamo il sistema:
$$\begin{cases} a = -1 \\ b = 1 \\ -5b = -3 \\ -a + 2b = 0 \end{cases}$$

che è evidentemente un sistema senza soluzioni. Dunque S non è un sottospazio vettoriale di $\mathbb{R}^4$ e il sistema che stiamo cercando è un sistema lineare non omogeneo in 4 variabili x, y, z, t:

$$A \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \alpha_n \end{pmatrix}$$

con $A \in \mathcal{M}_{n\times 4}(\mathbb{R})$ di rango 2: solo due equazioni saranno linearmente indipendenti.

Risolvere l'esercizio significa descrivere mediante un sistema di equazioni lineari l'insieme di tutti e soli gli elementi $(x, y, z, t) \in (-1, 1, -3, 0) + \langle(1, 0, 0, -1), (0, 1, -5, 2)\rangle$, vale a dire tutti e soli gli elementi (x, y, z, t) tali che $(x + 1, y - 1, z + 3, t) \in \langle(1, 0, 0, -1), (0, 1, -5, 2)\rangle$. Osserviamo che il vettore $(x + 1, y - 1, z + 3, t)$ appartiene al sottospazio $\langle(1, 0, 0, -1), (0, 1, -5, 2)\rangle$ se e solo se il rango della matrice

$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -5 & 2 \\ x+1 & y-1 & z+3 & t \end{pmatrix}$$

è uguale a due. Riducendo la matrice in forma a scala per righe otteniamo la matrice
$$\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -5 & 2 \\ 0 & 0 & z-2+5y & t+x+3-2y \end{pmatrix}$$

che ha rango 2 se e solo se $\begin{cases} 5y + z - 2 = 0 \\ x - 2y + t + 3 = 0 \end{cases}$ Il sistema ottenuto è quello che stavamo cercando.

(ii) L'insieme T è un sottospazio vettoriale di $\mathbb{R}^3$? Vediamo se il vettore $(2, 0, -1)$ appartiene al sottospazio $\langle(3, 5, 10), (18, 10, 14)\rangle$: $(2, 0, -1) = \alpha(3, 5, 10) + \beta(18, 10, 14) = (3\alpha + 18\beta, 5\alpha + 10\beta, 10\alpha + 14\beta)$. Risolvendo, si ottiene $\alpha = -\frac{1}{3}$, $\beta = \frac{1}{6}$. Dunque $T = \langle(3, 5, 10), (18, 10, 14)\rangle$ è un sottospazio di $\mathbb{R}^3$ di dimensione 2 e il sistema che stiamo cercando è dunque un sistema lineare omogeneo in 3 variabili di rango 1. A questo punto si procede come in (i).

Esercizio 6.4.6 Si considerino, al variare di $k \in \mathbb{R}$, i sottospazi vettoriali $S_k = \langle(k, 0, k), (1, k, 1), (1, 1, k)\rangle$ di $\mathbb{R}^3$. Determinare $\cap_{k\in\mathbb{R}} S_k$.

Svolgimento. Calcoliamo, innanzitutto, la dimensione dei sottospazi S_k al variare di k. Dopo aver disposto i generatori di S_k sulle righe di una matrice, basterà calcolare il numero di righe linearmente indipendenti della matrice ottenuta, cioè il rango della matrice:

$$\dim(S_k) = \operatorname{rg}\begin{pmatrix} 1 & 1 & k \\ 1 & k & 1 \\ k & 0 & k \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & k \\ 0 & 1-k & k-1 \\ 0 & -k & k-k^2 \end{pmatrix} =$$

$$= \operatorname{rg}\begin{pmatrix} 1 & 1 & k \\ 0 & 1 & k^2-1 \\ 0 & -k & k-k^2 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 1 & 1 & k \\ 0 & 1 & k^2-1 \\ 0 & 0 & k^3-k^2 \end{pmatrix}.$$

Avendo ridotto la matrice a scalini per righe, è facile calcolare il suo rango: per ogni $k \in \mathbb{R}$ tale che $k^3 - k^2 = k^2(k-1) \neq 0$, $\dim(S_k) = 3$, vale a dire, $S_k = \mathbb{R}^3$. Se $k = 1$ o $k = 0$, $\dim(S_k) = 2$. Si ha pertanto: $\cap_{k\in\mathbb{R}} S_k = \mathbb{R}^3 \cap S_0 \cap S_1 = S_0 \cap S_1$. Si tratta dunque di determinare l'intersezione dei sottospazi $S_0 = \langle(1, 1, 0), (0, 1, -1)\rangle$ e $S_1 = \langle(1, 1, 1), (0, 1, 0)\rangle$. Impostiamo un metodo generale di risoluzione di questo tipo di problema: dobbiamo stabilire se esistono elementi di S_1, ovvero vettori di $\mathbb{R}^3$ della forma $a(1, 1, 1) + b(0, 1, 0) = (a, a+b, a)$, che appartengono anche al sottospazio S_0, che siano, cioè, combinazioni lineari dei generatori di S_0. In altre parole, si tratta di stabilire se esistono valori di a e b in $\mathbb{R}$, tali che

$$\operatorname{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ a & a+b & a \end{pmatrix} = 2.$$

Riduciamo la matrice a scalini per righe:

$$\mathrm{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ a & a+b & a \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & b & a \end{pmatrix} = \mathrm{rg}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & a+b \end{pmatrix}.$$

Quindi il rango della matrice trovata è 2 se e solo se $b = -a$. I vettori che appartengono sia ad S_0 che ad S_1 sono pertanto i vettori di $\mathbb{R}^3$ della forma $(a, a+b, a)$, dove $b = -a$, cioè, i vettori della forma $(a, 0, a)$. Dunque: $\cap_{k\in\mathbb{R}} S_k = S_0 \cap S_1 = \langle(1, 0, 1)\rangle$.

6.5 Esercizi proposti

Esercizio 6.5.1 Studiare, al variare del parametro reale h, le soluzioni del sistema nelle incognite x, y, z

$$\begin{cases} 2x + (h+1)y - z = 3 \\ 3x + (2h-1)y - z = 3 \\ hx + 2hz = -h^2 \\ x + (h-2)y = 0 \\ (1+h)x + (h-2)y + 2hz = -h. \end{cases}$$

Esercizio 6.5.2 Sia dato il sistema nelle variabili x_1, x_2, x_3, x_4:

$$\begin{cases} x_1 - x_2 + 2x_4 = b \\ ax_1 + ax_2 - 2x_4 = c \\ -ax_2 + (a+1)x_4 = a. \end{cases}$$

Si discuta il sistema al variare di $a, b, c \in \mathbb{R}$.

Esercizio 6.5.3 Risolvere il sistema lineare

$$\begin{cases} 6y - 2z = 1 \\ x - 2z = 0 \\ 12x + 5y = 2. \end{cases}$$

Esercizio 6.5.4 Determinare un sistema lineare che abbia come insieme di soluzioni l'insieme $S = (2, 3, 0) + \langle(-1, -1, 1)\rangle$.

Esercizio 6.5.5 Si consideri l'endomorfismo L di $\mathbb{R}^3$ la cui matrice, rispetto alla base canonica di $\mathbb{R}^3$, sia:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 3 & 0 & 6 \\ 1 & 6 & 8 \end{pmatrix}.$$

Determinare $\ker A$, ImA, $\ker A^t$, ImA^t. Determinare inoltre, al variare di $k \in \mathbb{R}$, la controimmagine mediante L del vettore $(k, 3, 7)$.

Esercizio 6.5.6 Si considerino, al variare di $h \in \mathbb{R}$, i sottospazi vettoriali

$$T_h = \langle (1, h, 0, 0), (0, h, 0, 1), (1, -h, h, 1), (1, -1, 1, 1) \rangle$$

di $\mathbb{R}^4$. Determinare $\cap_{h \in \mathbb{R}} T_h$.

Lezione 7

Matrici

Abbiamo visto come le matrici $\mathcal{M}_{m,n}(\mathbb{R})$ rappresentino l'insieme delle applicazioni lineari tra due spazi V e W di dimensione rispettivamente n e m, una volta scelta una base per ciascuno dei due spazi (ricordiamo che la scelta delle basi corrisponde ad identificare ogni vettore di V con un elemento di $\mathbb{R}^n$ e ogni elemento di W con uno di $\mathbb{R}^m$).

Se componessimo due applicazioni lineari, la funzione composta sarebbe ancora lineare? E che cosa potremmo dire della matrice associata all'applicazione composta?

In questo paragrafo mostreremo che la composizione di due funzioni lineari è ancora un'applicazione lineare ed introdurremo il prodotto (righe per colonne) di matrici.

7.1 Prodotto righe per colonne

Problema. Dati tre spazi vettoriali V,W e Z, di dimensione, rispettivamente, n, m e p, per ciascuno di essi scegliamo una base: $\mathcal{V} = \{v_1, \ldots, v_n\}$, $\mathcal{W} = \{w_1, \ldots, w_m\}$ e $\mathcal{Z} = \{z_1, \ldots, z_p\}$. Siano poi $f \in \mathcal{L}in(V, W)$ e $g \in \mathcal{L}in(W, Z)$. Grazie alla scelta delle basi possiamo associare ad f una matrice $A_f \in \mathcal{M}_{m,n}(\mathbb{R})$ e a g una matrice $A_g \in \mathcal{M}_{p,m}(\mathbb{R})$.

1. La composizione $g \circ f$ è un elemento di $\mathcal{L}in(V, Z)$?

2. Se la risposta alla precedente domanda è affermativa, possiamo associare a $g \circ f$ la matrice $A_{g\circ f} \in \mathcal{M}_{p,n}(\mathbb{R})$. È possibile determinare la matrice $A_{g\circ f}$ conoscendo A_f e A_g?

Soluzione del problema. Per prima cosa dimostriamo che la composizione di due applicazioni lineari è ancora un'applicazione lineare. Siano quindi f e g due applicazioni lineari e sia $g \circ f : V \to Z$ la loro composizione. Per ogni $v \in V$ si ha $(g \circ f)(v) = g(f(v))$ con $f(v) \in W$. Ora, se prendiamo $v_1, v_2 \in V$,

$$\begin{aligned}(g \circ f)(v_1 +_V v_2) &= g(f(v_1 +_V v_2)) = g(f(v_1) +_W f(v_2)) = \\ &= g(f(v_1)) +_Z g(f(v_2)) = (g \circ f)(v_1) +_Z (g \circ f)(v_2)\end{aligned}$$

dove abbiamo usato la linearità di f e g. Sempre per linearità, se $\lambda \in \mathbb{R}$ e $v \in V$ allora $(g \circ f)(\lambda v) = g(f(\lambda v)) = g(\lambda f(v)) = \lambda g(f(v)) = \lambda(g \circ f)(v)$. Quindi la funzione composta $g \circ f$ è lineare.

Come calcolare la matrice ad essa associata rispetto alle basi $\mathcal{V}$ del dominio e $\mathcal{Z}$ del codominio? Sfruttiamo quello che sappiamo: le colonne di tale matrice sono le coordinate nella base $\mathcal{Z}$ delle immagini tramite $g \circ f$ dei vettori della base $\mathcal{V}$. Ad esempio, nella prima colonna dobbiamo mettere le coordinate del vettore $(g \circ f)(v_1)$, cioè $g(f(v_1))$, ma $f(v_1)$ è un vettore di W le cui coordinate nella base $\mathcal{W}$ sono le entrate della prima colonna di A_f. Quindi $g(f(v_1))$ è dato dal prodotto della matrice A_g per la prima colonna di A_f. Cioè, detta c_1 la prima colonna di A_f, la prima colonna di $A_{g \circ f}$ è $A_g c_1$, quindi una colonna formata da p righe. Si può (e si deve) fare lo stesso ragionamento per ogni vettore della base $\mathcal{V}$ e, quindi, per ogni colonna $c_1, c_2, \dots, c_n$ di A_f. Mettendo una dopo l'altra tali colonne (nell'ordine) formate da p-entrate si ottiene una matrice a p-righe e n-colonne, in cui ogni colonna rappresenta le coordinate dei vettori $(g \circ f)(v_i)$, $i = 1, \dots, n$, nella base $\mathcal{Z}$: è la matrice $A_{g \circ f}$ associata alla applicazione lineare $g \circ f$ rispetto alle basi $\mathcal{V}$ e $\mathcal{Z}$. Tale matrice è pertanto legata alle matrici A_f e A_g e si ha formalmente che, se

$$A_f = (a_{ij})_{i=1,\dots,m}^{j=1,\dots,n} \qquad A_g = (b_{ki})_{k=1,\dots,p}^{i=1,\dots,m},$$

allora

$$A_{g \circ f} = (c_{kj})_{k=1,\dots,p}^{j=1,\dots,n}$$

dove

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}.$$

In altre parole l'elemento che si trova nella riga k-esima e nella colonna j-esima di $A_{g \circ f}$ è il prodotto della riga k-esima di A_g per la colonna j-esima di A_f. In questo senso si chiama prodotto righe per colonne:

$$A_{g\circ f} = A_g A_f$$

con $A_f \in \mathcal{M}_{m,n}(\mathbb{R})$, $A_g \in \mathcal{M}_{p,m}(\mathbb{R})$ e $A_{g\circ f} \in \mathcal{M}_{p,n}(\mathbb{R})$.

Esempio 7.1.1 Consideriamo le due matrici

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Calcoliamo il prodotto AB. La matrice AB rappresenta, secondo l'interpretazione precedente e con una opportuna scelta delle basi, la composizione di una applicazione tra due spazi di dimensione 2 (nelle notazioni precedenti $B = A_f$) e di una applicazione da uno spazio di dimensione 3 ad uno di dimensione 2 ($A = A_g$). Moltiplicare A per B non è possibile! Lo si può vedere direttamente scrivendo le matrici

$$AB = \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}:$$

il numero di entrate di ogni riga di A (i.e. il numero di colonne di A) non è uguale al numero di entrate di ogni colonna di B (i.e. il numero di righe di B). Viceversa è possibile calcolare il prodotto BA. Questo corrisponde alla composizione di una funzione lineare da uno spazio di dimensione 3 ad uno spazio di dimensione 2 seguita da un endomorfismo di uno spazio di dimensione 2 (nelle notazioni precedenti questa volta $A_g = B$ e $A_f = A$). Tutto questo è formalmente compatibile. Si ha:

$$BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 5 \\ 2 & -1 & 7 \end{pmatrix}$$

dove, ad esempio, l'entrata di posto 2, 2 di BA è data dal "prodotto" della seconda riga di B per la seconda colonna di A: $1(1) + 2(-1) = -1$.

Esempio 7.1.2 Il nostro esempio mostra che, affinché il prodotto AB tra due matrici A e B sia definito, occorre che il numero di colonne di A sia eguale al numero di righe di B, cioè $A \in \mathcal{M}_{n_1,n_2}(\mathbb{R})$ e $B \in \mathcal{M}_{n_2,n_3}(\mathbb{R})$. Allora la matrice prodotto avrà tante righe quante quelle di A e tante colonne quante quelle di B cioè : $AB \in \mathcal{M}_{n_1,n_3}(\mathbb{R})$.

Notiamo che il prodotto righe per colonne di due matrici A e B era già stato introdotto nel paragrafo 5.3 nel caso particolare in cui $A \in \mathcal{M}_{m,n}(\mathbb{R})$ e $B \in \mathcal{M}_{n,1}(\mathbb{R})$. Per quanto appena osservato questo prodotto è ben definito e dà luogo ad una matrice $m \times 1$ cioè ad un vettore colonna di $\mathbb{R}^m$.

Osservazione 7.1.3 Dall'esempio precedente emerge con evidenza il fatto che il prodotto di matrici non è commutativo. Addirittura nell'esempio precedente il prodotto BA risultava definito, ma non risultava definito il prodotto AB. Facciamo un altro esempio: consideriamo le matrici

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Allora $AB = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ e $BA = \begin{pmatrix} 1 & 5 \\ 1 & 5 \end{pmatrix}$.

Osserviamo che, in modo equivalente, la composizione di applicazioni (lineari) non è un'operazione commutativa. Ad esempio, se consideriamo gli endomorfismi f e g di $\mathbb{R}^2$ di matrici A e B rispetto alla base canonica di $\mathbb{R}^2$, naturalmente $f \circ g \neq g \circ f$.

Osservazione 7.1.4 Il prodotto righe per colonne di matrici gode della proprietà associativa e della proprietà distributiva rispetto alla somma. Questo significa che, prese $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{p,m}(\mathbb{R})$ e infine $C \in \mathcal{M}_{q,p}(\mathbb{R})$: i prodotti $BA \in \mathcal{M}_{p,n}(\mathbb{R})$ e $CB \in \mathcal{M}_{q,m}(\mathbb{R})$ sono ben definiti, come pure i prodotti $C(BA) \in \mathcal{M}_{q,n}(\mathbb{R})$ e $(CB)A \in \mathcal{M}_{q,n}(\mathbb{R})$; si ha allora

$$C(BA) = (CB)A.$$

Per questo indicheremo questo prodotto semplicemente con CBA.

Inoltre, date $A_1 \in \mathcal{M}_{m,n}(\mathbb{R})$, $A_2 \in \mathcal{M}_{m,n}(\mathbb{R})$ e $B \in \mathcal{M}_{p,m}(\mathbb{R})$, allora $B(A_1 + A_2) = BA_1 + BA_2 \in \mathcal{M}_{p,n}(\mathbb{R})$ (distributività del prodotto rispetto alla somma).

Infine, per ogni $\alpha \in \mathbb{R}$, $B(\alpha A) = \alpha(BA)$.

Tutte le proprietà qui elencate possono essere dimostrate utilizzando la definizione di prodotto righe per colonne.

7.2 Matrici invertibili

Definizione 7.2.1 *Sia dato uno spazio vettoriale V di dimensione n. Definiamo lo spazio vettoriale degli endomorfismi di V, e lo denotiamo con $\mathcal{E}nd(V)$, come l'insieme delle applicazioni lineari di V in se stesso.*

Proposizione 7.2.2 *Se $f \in \mathcal{E}nd(V)$ è invertibile, l'applicazione inversa di f è anch'essa un'applicazione lineare che indicheremo con f^{-1}.*

Dimostrazione. Dobbiamo verificare che f^{-1} è lineare. Dati quindi $v_1, v_2 \in V$ dobbiamo vedere che $f^{-1}(v_1+v_2) = f^{-1}(v_1)+f^{-1}(v_2)$. Essendo f biiettiva esiste un unico $w_1 \in V$ tale che $f(w_1) = v_1$, così come esiste un unico $w_2 \in V$ tale che $f(w_2) = v_2$, cioè $f^{-1}(v_1) = w_1$ e $f^{-1}(v_2) = w_2$. In particolare allora $f(w_1 + w_2) = v_1 + v_2$ e quindi $f^{-1}(v_1 + v_2) = w_1 + w_2$, cioè $f^{-1}(v_1 + v_2) = f^{-1}(v_1) + f^{-1}(v_2)$. Ancora: dobbiamo mostrare che $f^{-1}(\lambda v) = \lambda f^{-1}(v)$ per ogni reale λ e ogni vettore $v \in V$. Sia dunque $w \in V$ il solo elemento di V tale che $f(w) = v$, cioè $f^{-1}(v) = w$. Per la linearità di f si ha che $f(\lambda w) = \lambda v$ e quindi, poiché la funzione è biiettiva, $f^{-1}(\lambda v) = \lambda w = \lambda f^{-1}(v)$. Concludiamo che f^{-1} è lineare. **C. V. D.**

Abbiamo dimostrato che se $f \in \mathcal{E}nd(V)$ è invertibile allora f^{-1} è un elemento di $\mathcal{E}nd(V)$. Scegliamo allora una base $\mathcal{V} = \{v_1, \dots, v_n\}$ del dominio di f e una base $\mathcal{W} = \{w_1, \dots, w_n\}$ del suo codominio, scegliamo cioè due basi di V (che magari possono coincidere). Allora se indichiamo con $A \in \mathcal{M}_n(\mathbb{R})$ la matrice associata a f rispetto alle basi $\mathcal{V}$ e $\mathcal{W}$, e se denotiamo con $B \in \mathcal{M}_n(\mathbb{R})$ la matrice di f^{-1} rispetto alla base $\mathcal{W}$ nel dominio e $\mathcal{V}$ nel codominio, allora il prodotto $BA \in \mathcal{M}_n(\mathbb{R})$ rappresenta la matrice della applicazione identica $f^{-1} \circ f = id_V$ rispetto alla base $\mathcal{V}$ sia nel dominio che nel codominio. Tale matrice avrà nella prima colonna le coordinate dell'immagine di v_1 rispetto alla applicazione identica, cioè v_1 stesso, nella base $\{v_1, \dots, v_n\}$. Le sue coordinate sono $(1, 0, 0, \dots, 0)$, così come l'immagine di v_2 è v_2 e le sue coordinate nella base $\mathcal{V}$ sono $(0, 1, 0, \dots, 0)$. La matrice ottenuta si chiama matrice *identica di ordine n* e si indica con I_n: è una matrice quadrata di ordine n in cui tutte le entrate sono nulle salvo quelle sulla diagonale che sono uguali a 1.

Per lo stesso motivo si ha che $AB = I_n$ rappresenta l'applicazione identica di V in se stesso espressa rispetto alla base $\mathcal{W}$ sia nel dominio che nel codominio.

Osservazione 7.2.3 Si noti che se $C \in \mathcal{M}_{n,p}(\mathbb{R})$ e $I_n \in \mathcal{M}_n(\mathbb{R})$ allora $I_n C = C$. (Attenzione: il prodotto CI_n non è definito! Se, però, I_p è la matrice identica di ordine p allora $CI_p = C$).

Definizione 7.2.4 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ si dice invertibile se esiste una matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che $AB = BA = I_n$. La matrice B si dice inversa di A e si denota con A^{-1}.*

Quali sono le proprietà delle matrici invertibili? Possiamo fare alcune brevi osservazioni sfruttando l'interpretazione di una matrice quadrata come applicazione lineare o, più precisamente, come endomorfismo. Sia $A \in \mathcal{M}_n(\mathbb{R})$ una matrice invertibile. Allora essa descrive un endomorfismo invertibile di $\mathbb{R}^n$. In particolare il rango di A è n e il numero di pivots in una qualunque forma a scalini per righe di A è n. In questa situazione non vi sono molte possibilità dal momento che la matrice ha ordine n: la forma a scalini deve avere tutti i pivots sulla diagonale e ogni elemento della diagonale deve essere un pivot. Questo fa sì che si possa ulteriormente raffinare la forma a scalini. Infatti, dopo opportune operazioni elementari sulle righe, la matrice sarà del tipo

$$\begin{pmatrix} \alpha_1 & * & * & * & \dots & \beta_1 \\ 0 & \alpha_2 & * & * & \dots & \beta_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \alpha_{n-1} & \beta_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}$$

dove $\alpha_i \neq 0$, e dove $*$ e β_j sono numeri reali. A questo punto possiamo continuare a semplificare la matrice con altri cambiamenti di base nel codominio. Ad esempio se sostituiamo la prima riga con la somma della prima riga con l'ultima moltiplicata per $-\frac{\beta_1}{\alpha_n}$ (cosa che può essere fatta poiché $\alpha_n \neq 0$) otterremo una matrice in cui l'elemento di posto $1, n$ (1 è la riga e n la colonna) è uguale a zero. Possiamo ripetere l'operazione per la seconda riga sommandole l'ultima moltiplicata per $-\frac{\beta_2}{\alpha_n}$. In questo modo si trova una matrice in cui gli elementi di posto k, n (k è la riga e n la colonna) con $k \neq n$ sono nulli e quello di posto n, n è α_n. Procedendo in questo modo con le altre colonne si arriva ad una matrice diagonale cioè ad una matrice del tipo:

$$\begin{pmatrix} \alpha_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \alpha_2 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \alpha_n \end{pmatrix}.$$

Possiamo fare ancora di meglio: se moltiplichiamo la prima riga per $\frac{1}{\alpha_1}$, la seconda per $\frac{1}{\alpha_2}$ e così via, alla fine otteniamo la matrice identica I_n.

Proposizione 7.2.5 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ è invertibile se e solo se mediante operazioni elementari sulle righe può essere ridotta alla matrice identica I_n.*

Dimostrazione. "⇒" Questa implicazione è stata appena vista: se una matrice è invertibile allora esiste una forma a scalini per righe uguale alla matrice identica. "⇐" Supponiamo che una matrice A abbia una forma a scalini per righe uguale alla matrice identica I_n. La riduzione in forma a scalini per righe è prodotta tramite operazioni elementari che non alterano il rango della matrice che quindi è uguale a n. Pertanto l'applicazione associata è invertibile e quindi la matrice è invertibile. **C.V.D.**

Osservazione 7.2.6 Se una matrice $A \in \mathcal{M}_n(\mathbb{R})$ ammette inversa, i.e., se esiste una matrice $B \in \mathcal{M}_n(\mathbb{R})$ tale che $AB = BA = I_n$, allora B è unica. Supponiamo infatti che esista un'altra matrice $C \in \mathcal{M}_n(\mathbb{R})$ tale che $CA = AC = I_n$. Consideriamo allora la matrice CAB. Essa è una matrice quadrata di ordine n e, per la proprietà associativa del prodotto righe per colonne, abbiamo $CAB = C(AB) = (CA)B$. Del resto $AB = CA = I_n$. Dunque abbiamo $CAB = CI_n = I_nB$, cioè $C = B$. Dunque l'inversa di una matrice è unica.

Per riconoscere se una matrice sia invertibile o meno non occorre ridurla alla matrice identica (bastava già una forma a scalini in cui tutti i pivots fossero gli elementi della diagonale). Tuttavia il lavoro fatto non è superfluo: ci fornisce senza ulteriore fatica un metodo per calcolare l'inversa di una matrice (invertibile)! Perché? Quello che abbiamo fatto è stato trovare un modo per ridurre la matrice di partenza A alla matrice identica. In termini molto grossolani è come se avessimo moltiplicato A^{-1} per A. Ma allora se applicassimo le stesse trasformazioni alla matrice identica I_n otterremmo il prodotto di A^{-1} per I_n cioè: $A^{-1}I_n = A^{-1}$. Quindi otterremmo la matrice inversa A^{-1}! Illustriamo quanto appena detto attraverso un esempio.

Esempio 7.2.7 Prendiamo la matrice

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

cerchiamo di vedere se è invertibile ed eventualmente calcoliamone l'inversa. Procediamo come abbiamo detto prima e consideriamo la matrice che si ottiene formalmente scrivendo le colonne della matrice I_3 a destra di A in un'unica nuova matrice:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

Ora applichiamo a tutta la matrice esattamente le stesse trasformazioni per riga che useremmo per ottenere una forma a scalini per righe, diagonale, della matrice A. Cominciamo: scambiamo l'ultima riga di A con la prima:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 & 0 \end{array}\right); \tag{7.1}$$

sostituiamo alla terza riga la somma della terza riga con la prima moltiplicata per 2:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 2 \end{array}\right);$$

sostituiamo alla terza riga la somma della terza riga con la seconda moltiplicata per -3:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array}\right);$$

sostituiamo alla prima riga la somma della prima con l'ultima moltiplicata per $-\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} -1 & 0 & 0 & -\frac{1}{3} & 1 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 & -3 & 2 \end{array}\right);$$

moltiplichiamo la prima riga per -1 e l'ultima per $\frac{1}{3}$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{3} & -1 & \frac{2}{3} \end{array}\right).$$

La matrice

$$\begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix}$$

è dunque la matrice inversa della matrice di partenza. (N.B. Consideriamo 7.1: se moltiplichiamo la matrice A per la matrice che è formata dalle ultime 3 colonne di 7.1, abbiamo

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix},$$

cioè la matrice data dalle prime tre colonne di 7.1. Questo significa che la matrice data dalle ultime tre colonne di 7.1 è la matrice che moltiplicata per A dà la matrice data dalle prime tre colonne di 7.1. La stessa considerazione può essere fatta per le successive trasformazioni per riga: di fatto tutte le trasformazioni di righe e colonne possono essere realizzate tramite moltiplicazione per opportune matrici invertibili). Possiamo verificare direttamente, usando il prodotto righe per colonne, che la matrice trovata è effettivamente la matrice inversa della matrice di partenza:

$$\begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Esempio 7.2.8 Consideriamo il seguente sistema nelle incognite x_1, x_2, x_3:

$$A\mathbf{x} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}.$$

La matrice $A \in \mathcal{M}_3(\mathbb{R})$ è quella dell'esempio precedente e quindi è invertibile. Il sistema avrà dunque un'unica soluzione data dall'unico vettore di $\mathbb{R}^3$ che è spedito, dall'applicazione rappresentata da A, nel vettore $(1, -2, 0)$, o, equivalentemente, dall'unico vettore immagine di $(1, -2, 0)$ mediante l'applicazione inversa. Poiché conosciamo la matrice dell'applicazione inversa siamo in grado di risolvere il sistema:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -1 & -\frac{1}{3} \\ 0 & 1 & 0 \\ \frac{1}{3} & -1 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ -2 \\ \frac{7}{3} \end{pmatrix}.$$

Esercizio 7.2.9 *(***Problema A**) A questo punto siamo in grado di affrontare il problema indicato nell'introduzione come problema A. Esso si può configurare come un sistema in due incognite E_1 e E_2 dato da

$$\begin{cases} E_1 + 2E_2 = 22 \\ 3E_1 - 2E_2 = 2 \end{cases}.$$

Il rango della matrice incompleta è due e quindi uguale al rango della matrice completa. Il sistema è risolubile. Le soluzioni sono date da una soluzione particolare sommata al nucleo dell'applicazione lineare associata alla matrice del sistema incompleto. Ma tale matrice è la matrice di una applicazione lineare tra due spazi vettoriali di dimensione 2. Poiché la matrice ha rango 2 l'applicazione è suriettiva e quindi è un isomorfismo. Il nucleo è pertanto banale. Il sistema ha quindi una e una sola soluzione: quella che abbiamo trovato nella introduzione ($E_1 = 6, E_2 = 8$)!

Per risolvere la seconda parte del problema usiamo tre incognite E_1, E_2 e Z. Il sistema associato al problema sarà un sistema di due equazioni in tre incognite:

$$\begin{cases} E_1 - E_2 - Z = 2 \\ -E_1 + E_2 + 2Z = 4. \end{cases}$$

La matrice completa del sistema è data da

$$\left(\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ -1 & 1 & 2 & 4 \end{array}\right)$$

e si vede subito che il rango della matrice completa e il rango della matrice incompleta sono uguali a 2. Il sistema ammette soluzioni.

Occorre determinare il nucleo della matrice incompleta

$$\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 2 \end{pmatrix}$$

che è la matrice di una applicazione lineare da $\mathbb{R}^3$ in $\mathbb{R}^2$, di rango due. Per la formula delle dimensioni il nucleo di questa applicazione ha dimensione 1 ed è dato da $\langle(1, 1, 0)\rangle$.

Una soluzione particolare del sistema è fornita già nella introduzione: $E_1 = 12$, $E_2 = 4$, $Z = 6$. Le soluzioni sono allora tutte e sole della forma: $E_1 = 12 + t$, $E_2 = 4 + t$, $Z = 6$ al variare di $t \in \mathbb{R}$.

7.3 Esercizi svolti

Esercizio 7.3.1 Calcolare i prodotti AB e BA delle matrici

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -3 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 1 \\ -2 & 1 \\ -1 & 0 \\ 3 & 5 \end{pmatrix}.$$

Indicato con $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'endomorfismo di $\mathbb{R}^2$ la cui matrice rispetto alla base canonica di $\mathbb{R}^2$ sia la matrice AB, determinare $f(2, -3)$. La matrice BA è la matrice di un endomorfismo?

Svolgimento. Per calcolare AB e BA dobbiamo semplicemente usare la definizione di prodotto righe per colonne:

$$AB = \begin{pmatrix} 5 & 23 \\ 7 & 11 \end{pmatrix}; \qquad BA = \begin{pmatrix} 0 & 1 & -3 & 2 \\ -2 & -3 & -9 & -6 \\ -1 & -2 & -3 & -4 \\ 3 & 11 & -6 & 22 \end{pmatrix}.$$

Dunque, AB è una matrice quadrata di ordine 2 e quindi la matrice associata ad un endomorfismo di $\mathbb{R}^2$. BA è una matrice quadrata di ordine 4 ed è pertanto la matrice associata ad un endomorfismo di $\mathbb{R}^4$. Dire che AB è la matrice associata ad un endomorfismo f rispetto alla base canonica significa che le colonne di f sono le coordinate rispetto alla base canonica e_1, e_2 di $\mathbb{R}^2$ dei vettori $f(e_1)$ e $f(e_2)$. Così $f(e_1) = 5e_1 + 7e_2 = (5, 7)$ e $f(e_2) = 23e_1 + 11e_2 = (23, 11)$. Usiamo la linearità di f per calcolare l'immagine del vettore $(2, -3)$:

$f(2, -3) = f(2e_1 - 3e_2) = 2f(e_1) - 3f(e_2) = 2(5, 7) - 3(23, 11) = (-59, -19)$.

Esercizio 7.3.2 Determinare la matrice inversa di ciascuna delle seguenti matrici:

$$M_1 = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3. \end{pmatrix}$$

Svolgimento. Procediamo come illustrato nell'esempio 7.2.7. Cominciamo dalla matrice M_1 e consideriamo la matrice $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{array}\right)$.

Sostituiamo alla seconda riga la somma della prima riga moltiplicata per due e della seconda: $\left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 1 \end{array}\right)$.

Ora sostituiamo alla prima riga la somma della prima riga moltiplicata per -3 e della seconda: $\left(\begin{array}{cc|cc} -3 & 0 & -1 & 1 \\ 0 & 3 & 2 & 1 \end{array}\right)$.

Infine, moltiplichiamo la prima riga per $-\frac{1}{3}$ e la seconda per $\frac{1}{3}$:
$\left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & \frac{2}{3} & \frac{1}{3} \end{array}\right)$.
Abbiamo così ottenuto la matrice inversa della matrice M_1:

$$M_1 = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Analogamente si procede per la matrice M_2:
$\left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 4 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 2 & 0 & 0 & -2 & 0 & 1 \\ 0 & -1 & 0 & 0 & 1 & -2 \\ 0 & 0 & 2 & 0 & 0 & 1 \end{array}\right) \Rightarrow$
$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 0 & 0 & \frac{1}{2} \end{array}\right)$.
Dunque $M_2^{-1} = \begin{pmatrix} -1 & 0 & \frac{1}{2} \\ 0 & -1 & 2 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$. Osserviamo che M_2 è una matrice triangolare superiore e che la sua inversa è anch'essa una matrice triangolare superiore. Questo fatto è vero in generale.

Anche nel caso della matrice M_3 si procede come per qualsiasi altra matrice, ma, essendo essa una matrice diagonale, il numero di passaggi sarà inferiore: la forma diagonale è infatti già di per sé una forma a scalini per righe. Otteniamo dunque, immediatamente,

$$M_3^{-1} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}.$$

In generale, data una matrice diagonale di ordine n con elementi diagonali non nulli α_1, α_2, $\dots, \alpha_n$ (se uno di questi fosse nullo la matrice non sarebbe invertibile!), la sua inversa è la matrice diagonale con elementi diagonali α_1^{-1}, α_2^{-1}, $\dots, \alpha_n^{-1}$.

Esercizio 7.3.3 Sia $h : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare associata, rispetto alla base canonica di $\mathbb{R}^3$, alla matrice:

$$H = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & -2 \\ 0 & -1 & -3 \end{pmatrix}.$$

Stabilire se h è ivertibile e determinare $h^{-1}(2,1,2)$.

Svolgimento. Si verifica facilmente che il rango della matrice H è uguale a 3, pertanto la funzione h è invertibile. A questo punto un modo di risolvere l'esercizio è calacolare la matrice inversa della matrice H. Infatti se h è un endomorfismo invertibile di uno spazio vettoriale V e H è la matrice ad esso associata rispetto ad una base $\mathcal{V}$ del dominio e ad una base $\mathcal{W}$ del codominio, allora la matrice associata all'endomorfismo inverso h^{-1}, rispetto alla base $\mathcal{W}$ del dominio e alla base $\mathcal{V}$ del codominio, è la matrice H^{-1}.

Procediamo col calcolo della matrice H^{-1}:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & -1 & -3 & 0 & 0 & 1 \end{array}\right) \Rightarrow$$
$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow$$
$$\left(\begin{array}{ccc|ccc} \frac{33}{2} & 0 & 0 & \frac{15}{2} & -9 & 6 \\ 0 & -33 & 0 & -9 & -9 & 6 \\ 0 & 0 & -11 & 1 & 1 & 3 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ 0 & 1 & 0 & \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ 0 & 0 & 1 & -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{array}\right).$$

Abbiamo quindi

$$H^{-1} = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix}.$$

Pertanto $h^{-1}(2,1,2) = \begin{pmatrix} \frac{5}{11} & -\frac{6}{11} & \frac{4}{11} \\ \frac{3}{11} & \frac{3}{11} & -\frac{2}{11} \\ -\frac{1}{11} & -\frac{1}{11} & -\frac{3}{11} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{12}{11} \\ \frac{5}{11} \\ -\frac{9}{11} \end{pmatrix}.$

7.4 Esercizi proposti

Esercizio 7.4.1 Calcolare l'inversa della matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

Esercizio 7.4.2 Calcolare $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$.

Esercizio 7.4.3 Sia $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Costruire, se possibile, una matrice $B \neq I_2$ tale che $(A+B)(A-B) = A^2 - B^2$.

Esercizio 7.4.4 Una matrice quadrata A si dice *nilpotente* se $A^k = 0$ per qualche intero $k > 0$. Mostrare che se A è nilpotente allora $I + A$ è invertibile.

Esercizio 7.4.5 Data la matrice $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \\ 2 & 5 \end{pmatrix}$,

1. trovare, se possibile, una matrice B tale che $BA = I_2$. Una siffatta matrice B è unica?
2. Trovare, se possibile, una matrice C tale che $AC = I_3$.

Lezione 8

Determinante, cambiamenti di base

Il fatto di riconoscere se una matrice quadrata sia invertibile o meno, al di là di dirci se l'applicazione da essa rappresentata sia un isomorfismo, ha ripercussioni sul calcolo del rango di una matrice qualsiasi. Vediamo perché.

8.1 Minori

Definizione 8.1.1 *Data una matrice $A \in \mathcal{M}_{m,n}(\mathbb{R})$, si dice minore di A una matrice che si ottiene da A eliminando alcune sue righe e colonne.*

Ad esempio, la matrice $(2) \in \mathcal{M}_{1,1}(\mathbb{R})$ è una sottomatrice di $B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 21 \end{pmatrix}$. Si ottiene infatti eliminando la seconda riga, la prima e la terza colonna di B. Un minore si dice quadrato se è una matrice quadrata. In una matrice che appartiene allo spazio vettoriale $\mathcal{M}_{m,n}(\mathbb{R})$, con $m \leq n$, si possono solamente trovare minori quadrati di ordine compreso tra 1 e m. Ad esempio una matrice $B \in \mathcal{M}_{2,3}(\mathbb{R})$ possiede esattamente 3 minori quadrati di ordine 2!

Sappiamo che una matrice quadrata di ordine n ha rango massimo se e solo se è invertibile. Possiamo allora caratterizzare come segue il rango di una matrice:

Proposizione 8.1.2 *Sia A una matrice in $\mathcal{M}_{m,n}(\mathbb{R})$. Allora il rango di A è uguale al massimo ordine s di un suo minore quadrato invertibile.*

Dimostrazione. Sia $p = \mathrm{rg}A$. Allora $p \leq m$ e $p \leq n$, perché p è il massimo numero di colonne linearmente indipendenti di A ed è uguale al massimo numero di righe linearmente indipendenti di A. Questo vuol dire che esistono p righe di A linearmente indipendenti. Togliamo da A le altre righe ottenendo così un minore $p \times n$ di A, che indicheremo con A'. Se A' fosse quadrato sarebbe invertibile, avendo $p = n$ righe linearmente indipendenti e quindi rango massimo. Altrimenti sappiamo che A' ha rango p e quindi possiede pure p colonne linearmente indipendenti. Togliamo da A' le colonne che non sono linearmente indipendenti: otteniamo una sottomatrice quadrata di A di ordine p e di rango massimo p. Così $p = \mathrm{rg}A \leq s$, per definizione di massimo.

Sia ora A' un minore quadrato di A, invertibile e di massimo ordine s con questa proprietà. Dire che è invertibile vuol dire che le righe di A' sono tutte linearmente indipendenti: ma le righe di tale minore sono ottenute prendendo delle righe di A ed eliminando alcune colonne. Queste righe di A devono essere vettori di $\mathbb{R}^n$ linearmente indipendenti perché se fossero linearmente dipendenti allora i vettori ottenuti eliminando qualche loro entrata (i.e. qualche colonna di A) sarebbero pure linearmente dipendenti! Così $s \leq \mathrm{rg}A$. Abbiamo provato che $\mathrm{rg}A \leq s \leq \mathrm{rg}A$. Vale pertanto l'uguaglianza. **C.V.D.**

Dalla proposizione precedente si deduce che una matrice in $\mathcal{M}_{m,n}(\mathbb{R})$ ha come massimo rango possibile il valore minore tra n e m: in effetti tale numero è l'ordine del massimo minore quadrato che si può costruire. La proposizione precedente evidenzia inoltre l'importanza di saper riconoscere direttamente se una matrice quadrata sia invertibile o no. Il metodo per farlo sarà argomento della prossima sezione.

8.2 Il determinante

Intendiamo definire una funzione f avente come dominio le matrici quadrate, diciamo di ordine n, e come codominio i numeri reali, in modo che l'immagine $f(A)$ di una matrice quadrata A sia diversa da 0 se e solo se A è invertibile. Tale funzione esiste e si chiama *determinante*. Non vogliamo qui esporre la teoria completa di tale funzione, ma limitarci a dare alcune giustificazioni per la sua costruzione.

Caso n=1. L'insieme delle matrici reali 1×1 è $\mathcal{M}_1(\mathbb{R}) = \{(a) \mid a \in \mathbb{R}\}$, quindi il rango di una matrice quadrata di ordine 1, $A = (a)$, è massimo (e uguale ad 1) se e solo se $a \neq 0$. È facile costruire in questo caso l'applicazione determinante:

$$\begin{array}{rccc} \det: & \mathcal{M}_1(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ & A=(a) & \longmapsto & a. \end{array}$$

In questo modo una matrice $A = (a)$ è invertibile se e solo se $\det(A) \neq 0$.

Caso n=2. Cerchiamo ora di capire quando una matrice 2×2:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

è invertibile o, equivalentemente, quando non lo è, cioè quando non ha rango massimo. Se non ha rango massimo, allora vuol dire che le sue righe sono linearmente dipendenti. Cioè

$$\alpha(a, b) + \beta(c, d) = (0, 0)$$

per qualche coppia di numeri reali $(\alpha, \beta) \neq (0, 0)$. Sia $\alpha \neq 0$ (il caso $\beta \neq 0$ porterebbe alle stesse conclusioni). Abbiamo $(a, b) = -\frac{\beta}{\alpha}(c, d)$, cioè $a = -\frac{\beta}{\alpha}c$ e $b = -\frac{\beta}{\alpha}d$, dunque $ad = cb$ perciò:

$$ad - cb = 0.$$

Questa è la condizione affinché le righe di A siano linearmente dipendenti! D'altra parte se $ad - cb \neq 0$ allora le righe di A sono linearmente indipendenti e il rango di A è massimo. Siamo dunque in grado di costruire l'applicazione determinante per le matrici quadrate di ordine 2:

$$\begin{array}{rcl} \det: \mathcal{M}_2(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & \det(A) = ad - bc. \end{array}$$

Così $A \in \mathcal{M}_2(\mathbb{R})$ è invertibile se e solo se $\det A \neq 0$.

Torniamo al caso generale. Quali proprietà richiediamo alla funzione determinante?

i) Innanzitutto dovrà essere una funzione definita su $\mathcal{M}_n(\mathbb{R})$ a valori nell'insieme dei numeri reali:

$$\det \; : \; \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$$

tale che alla matrice A venga associato il numero reale $\det(A)$.

ii) La matrice identica $I_n \in \mathcal{M}_n(\mathbb{R})$ è invertibile e quindi richiederemo che abbia determinante non nullo. In particolare chiediamo che $\det I_n = 1$.

iii) Se una matrice ha due righe uguali, oppure due colonne uguali, oppure, più in generale, se il rango di una matrice $A \in \mathcal{M}_n(\mathbb{R})$ non è massimo, (cioè è minore di n), allora $\det A = 0$ (in tutti questi casi infatti la matrice A non è invertibile!). In particolare $\det(\mathbf{0}_n) = 0$ ($\mathbf{0}_n$ è la matrice quadrata di ordine n nulla).

iv) NON possiamo aspettarci che l'applicazione determinante sia lineare. In effetti, già nel caso 2×2 in generale $\det(A + B) \neq \det A + \det B$. Ad esempio, prendiamo $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Allora

$$A + B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

e si ha: $\det(A + B) = 0$ mentre $\det(A) + \det(B) = 1 + 0 = 1$ (si noti che B non ha rango massimo).

v) La composizione di due isomorfismi è ancora un isomorfismo, cioè un'applicazione invertibile! Vorremo allora che $\det(AB) \neq 0$ se $\det A \neq 0$ e $\det B \neq 0$. Posssiamo dire anche qualcosa in più: se A rappresenta un isomorfismo di $\mathbb{R}^n$ e A^{-1} è la matrice inversa della matrice A, cioè la matrice che rappresenta l'endomorfismo inverso rispetto alle stesse basi, allora la matrice dell'endomorfismo composto è la matrice identica I_n. Vorremo pertanto che: $1 = \det I_n = \det(AA^{-1}) = \det A \;\det A^{-1}$, cioè $\det A^{-1} = \frac{1}{\det A}$. Più in generale ci aspettiamo che valga $\det(AB) = \det A \det B$.

A colpo d'occhio sembrerebbe assai complicato trovare una funzione che goda di tutte queste proprietà ma in effetti una funzione con tali caratteristiche esiste e può essere definita, a partire dagli esempi che abbiamo illustrato, procedendo per induzione sull'ordine delle matrici (nel senso che usando il determinante delle matrici 2×2 possiamo costruire il determinante delle matrici 3×3 e utilizzando il determinante delle matrici 3×3 possiamo costruire il determinante delle matrici 4×4 e così via...). Si verifichi che la funzione determinante definita sopra su $\mathcal{M}_1(\mathbb{R})$ e $\mathcal{M}_2(\mathbb{R})$ verifica le proprietà $i) - v)$.

Costruiamo ora la funzione determinante per qualunque matrice quadrata di ordine n. Abbiamo già affrontato i casi $n = 1$, $n = 2$. Ora supponiamo di essere in grado di calcolare il determinante di una matrice quadrata di ordine $(n - 1)$ e calcoliamo il determinante di una matrice quadrata A di ordine n.

Osserviamo preliminarmente che cancellando la riga $i-esima$ e la colonna $j-esima$ della matrice A si ottiene un minore quadrato di ordine $n-1$ di A che indichiamo con $\mathcal{A}_{ij}$. Per ipotesi (induttiva) conosciamo allora $\det(\mathcal{A}_{ij})$, per ogni $i, j = 1, \ldots, n$.

Per ogni elemento $a_{k1}, a_{k2}, \ldots, a_{kn}$ della riga $k-esima$ della matrice $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ possiamo considerare il minore associato, nell'ordine: $\mathcal{A}_{k1}, \mathcal{A}_{k2}, \ldots, \mathcal{A}_{kn}$. Diamo allora la seguente definizione:

Definizione 8.2.1 *Data la matrice* $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$, *poniamo*

$$\det(A) = a_{11} \qquad se \quad n = 1$$

e

$$\begin{aligned} \det A = & \det(a_{ij}) = (-1)^{k+1} a_{k1} \det(\mathcal{A}_{k1}) + (-1)^{k+2} a_{k2} \det(\mathcal{A}_{k2}) + \\ + & (-1)^{k+3} a_{k3} \det(\mathcal{A}_{k3}) + \cdots + (-1)^{k+n} a_{kn} \det(\mathcal{A}_{kn}) \end{aligned} \tag{8.1}$$

se $n > 1$, *con* $1 \leq k \leq n$.

Attenzione: l'espressione (8.1) si chiama sviluppo del determinante secondo la riga k-esima. Ogni riga può essere scelta per il calcolo del determinante: il risultato della formula (8.1) sarà sempre lo stesso! Vale ancora di più: se avessimo scelto una qualsiasi colonna avremmo potuto costruire una formula analoga e con lo stesso risultato! Vale cioè anche la seguente formula:

$$\begin{aligned} \det A = \det(a_{ij}) = & (-1)^{k+1} a_{1k} \det(\mathcal{A}_{1k}) + (-1)^{k+2} a_{2k} \det(\mathcal{A}_{2k}) + \\ + & (-1)^{k+3} a_{3k} \det(\mathcal{A}_{3k}) + \cdots + (-1)^{k+n} a_{nk} \det(\mathcal{A}_{nk}) \end{aligned}$$

(sviluppo del determinante di A rispetto alla colonna $k-esima$: gli elementi della colonna k-esima sono infatti $a_{1k}, a_{2k}, a_{3k}, \ldots, a_{nk}$).

Si osservi che la regola precedente può essere utilizzata per calcolare il determinante delle matrici 2×2. Il risultato sarà esattamente quello già descritto!

È sorprendente che una definizione così complicata possa implicare tutte le proprietà da noi richieste, ma in effetti questo è ciò che succede. La dimostrazione di questo fatto deriva da una teoria più avanzata di cui non ci occuperemo.

Esempi 8.2.2 *a*) Sappiamo già calcolare il determinante delle matrici di ordine 2. Utilizziamo la nostra formula nel caso di una matrice $B = (b_{ij}) \in$

$\mathcal{M}_3(\mathbb{R})$. Sia ad esempio

$$B = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}.$$

Sviluppiamo il determinante rispetto alla seconda riga ($b_{21} = 0, b_{22} = 3, b_{23} = 2$):

$$\det B = (-1)^{2+1}(0) \det \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} + (-1)^{2+2}(3) \det \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} +$$

$$+(-1)^{2+3}(2) \det \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = 3(2-(-1)) - 2(2) = 5.$$

Potremmo adesso sviluppare il determinante di B secondo la terza colonna ($b_{13} = -1, b_{23} = 2, b_{33} = 1$), ottenendo:

$$\det B = (-1)^{1+3}(-1) \det \begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix} + (-1)^{2+3}(2) \det \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} +$$

$$+(-1)^{3+3}(1) \det \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = (-1)(-3) + (-2)(2) + (6) = 3 - 4 + 6 = 5.$$

Come ci aspettavamo abbiamo trovato in entrambi i casi lo stesso risultato: i due modi di procedere sono equivalenti. Nello stesso modo avremmo potuto scegliere qualsiasi altra riga o colonna di B e avremmo trovato $\det(B) = 5$.

b) Come si è visto, se $a_{ij} = 0$, nello sviluppo del determinante rispetto alla i-esima riga o alla j-esima colonna l'addendo $(-1)^{i+j} a_{ij} \mathcal{A}_{ij}$ non dà alcun contributo. Pertanto, essendo liberi di scegliere la riga o la colonna rispetto a cui sviluppare, sceglieremo, se possibile, una riga o colonna con molte entrate nulle!

c) Se $A \in \mathcal{M}_n(\mathbb{R})$ è del tipo:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ 0 & & & & \\ 0 & & & & \\ \vdots & & \mathcal{A}_{11} & & \\ 0 & & & & \\ 0 & & & & \end{pmatrix}$$

cioè $a_{k1} = 0$ per ogni $k \neq 1$, allora $\det A = a_{11} \det \mathcal{A}_{11}$ (basta sviluppare secondo gli elementi della prima colonna). Analogamente, se

$$A = \begin{pmatrix} a_{11} & 0 & \dots & \dots & 0 \\ a_{21} & & & & \\ a_{31} & & & & \\ \vdots & & \mathcal{A}_{11} & & \\ a_{n-11} & & & & \\ a_{n1} & & & & \end{pmatrix},$$

$\det A = a_{11} \det \mathcal{A}_{11}$ (sviluppo secondo la prima riga).

Osservazione 8.2.3 Sappiamo che una matrice quadrata triangolare superiore con tutti gli elementi sulla diagonale diversi da zero ha rango massimo quindi è invertibile e ha determinante diverso da zero. Più precisamente, il determinante è il prodotto degli elementi sulla diagonale. Infatti se la matrice $A \in \mathcal{M}_n(\mathbb{R})$ è del tipo:

$$A = \begin{pmatrix} a_1 & a_{12} & \dots & & & a_{1n} \\ 0 & a_2 & a_{23} & \dots & & \vdots \\ 0 & 0 & a_3 & \dots & & \\ \vdots & \vdots & 0 & \ddots & & \\ 0 & \vdots & & & \ddots & \vdots \\ 0 & 0 & \dots & & & a_n \end{pmatrix}$$

allora possiamo sviluppare il determinante rispetto alla prima colonna:

$$\det A = a_1 \det \begin{pmatrix} a_2 & a_{12} & \dots & \\ 0 & a_3 & \dots & \\ 0 & 0 & \ddots & \\ 0 & & & \\ 0 & & & a_n \end{pmatrix}$$

e poi procedere nello stesso modo. Alla fine si ottiene che il determinante è il prodotto degli elementi sulla diagonale di A.

8.3 Calcolo dell'inversa

In questo paragrafo vogliamo mostrare come il calcolo del determinante ci fornisca un nuovo metodo per calcolare l'inversa di una matrice! Diamo innanzitutto una definizione che ci sarà utile in seguito.

Definizione 8.3.1 *Data una matrice $A \in \mathcal{M}_{n,m}(\mathbb{R})$ diremo trasposta di A, e la indicheremo con A^t, la matrice in $\mathcal{M}_{m,n}(\mathbb{R})$ che si ottiene da A scambiando le righe con le colonne.*

Esempio 8.3.2 La matrice trasposta di $\underline{\mathbf{x}} = (x_1\ x_2\ \dots\ x_n) \in \mathcal{M}_{1,n}(\mathbb{R})$ è la matrice $(\underline{\mathbf{x}})^t = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{R})$. La matrice trasposta della matrice $A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & -2 & 1 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{R})$ è la matrice $A^t = \begin{pmatrix} 2 & 0 \\ 3 & -2 \\ 4 & 1 \end{pmatrix}$.

Volendo essere più precisi potremmo dare le seguente definizione: se $A = (a_{ij}) \in \mathcal{M}_{n,m}(\mathbb{R})$, allora $A^t = (b_{kl}) \in \mathcal{M}_{m,n}(\mathbb{R})$ con $b_{kl} = a_{lk}$ (si noti che tutti gli indici sono ben definiti!).

Date due matrici qualsiasi $A \in \mathcal{M}_{n,m}(\mathbb{R})$ e $B \in \mathcal{M}_{m,p}(\mathbb{R})$ vale $(AB)^t = B^t A^t$. Provarlo per esercizio.

Come già preannunciato, lo scopo di questo paragrafo sarà fornire un metodo per trovare l'inversa di una matrice usando la nozione di determinante. A quel punto il determinante non sarà più solo un segnalatore di invertibilità!

Sia quindi $A \in \mathcal{M}_n(\mathbb{R})$ e supponiamo che $\det A \neq 0$ cioè che la matrice sia invertibile. Costruiamo A^{-1} procedendo per passi successivi:

- primo passo: per ogni posizione ij consideriamo il minore $\mathcal{A}_{ij}$ (di ordine $n-1$): calcoliamo $\det \mathcal{A}_{ij}$ e moltiplichiamolo per il segno associato $(-1)^{i+j}$, otteniamo un numero reale (che può essere anche zero): $(-1)^{i+j} \det \mathcal{A}_{ij}$ che indichiamo con $\mathcal{S}_{ij}$.

Osservazione : il determinante di A, sviluppato rispetto alla $i-esima$ riga, è: $\det A = a_{i1}\mathcal{S}_{i1} + a_{i2}\mathcal{S}_{i2} + \ldots + a_{in}\mathcal{S}_{in}$.

Costruiamo la matrice $\mathcal{S} = (\mathcal{S}_{ij})$;

- **secondo passo**: trasponiamo la precedente matrice, ottenendo $\mathcal{S}^t$. Per definizione di matrice trasposta le colonne di $\mathcal{S}^t$ sono le righe di $\mathcal{S}$;

- terzo passo: moltiplichiamo la matrice $\mathcal{S}^t$ per lo scalare $\frac{1}{\det A}$. Questo significa moltiplicare ogni entrata di $\mathcal{S}^t$ per $\frac{1}{\det A}$.

Abbiamo finito: la matrice trovata è l'inversa della matrice A! Perché? In effetti dobbiamo dimostrarlo. Calcoliamo allora il prodotto righe per colonne

$$\left(\frac{1}{\det A}\mathcal{S}^t\right) A.$$

Dal momento che moltiplicare la matrice $\mathcal{S}^t$ per $\frac{1}{\det A}$ significa moltiplicare ogni entrata di $\mathcal{S}^t$ per lo scalare $\frac{1}{\det A}$, abbiamo

$$\left(\frac{1}{\det A}\mathcal{S}^t\right) A = \frac{1}{\det A}(\mathcal{S}^t A).$$

Chi è l'elemento di posto $1,1$ della matrice $\mathcal{S}^t A$? È il prodotto della prima riga di $\mathcal{S}^t$ per la prima colonna di A. La prima riga di $\mathcal{S}^t$ è la prima colonna di $\mathcal{S}$ quindi la n-upla $(\mathcal{S}_{11}, \mathcal{S}_{21}, \mathcal{S}_{31}, \ldots, \mathcal{S}_{n1})$ che dobbiamo moltiplicare per la prima colonna di A. Otteniamo:

$$a_{11}\mathcal{S}_{11} + a_{21}\mathcal{S}_{21} + a_{31}\mathcal{S}_{31} + \cdots + a_{n1}\mathcal{S}_{n1}$$

ma questo è proprio $\det A$, più precisamente è lo sviluppo del determinante di A secondo la prima colonna di A! (I segni sono contenuti nella definizione degli $\mathcal{S}_{k1}$). Analogamente, l'elemento di posto i, i di $\mathcal{S}^t A$ è, per ogni $i = 1, \ldots, n$, il prodotto della riga i-esima di $\mathcal{S}^t$, i.e., della colonna i-esima di $\mathcal{S}$, per la colonna i-esima di A, vale a dire la somma:

$$a_{1i}\mathcal{S}_{1i} + a_{2i}\mathcal{S}_{2i} + a_{3i}\mathcal{S}_{3i} + \cdots + a_{ni}\mathcal{S}_{ni}.$$

Tale somma coincide con lo sviluppo del determinante di A secondo gli elementi della colonna i-esima e quindi con $\det A$.

Procediamo con l'elemento di posto $1,2$. In questo caso dobbiamo moltiplicare la prima riga di $\mathcal{S}^t$, che è sempre $(\mathcal{S}_{11}, \mathcal{S}_{21}, \mathcal{S}_{31}, \ldots, \mathcal{S}_{n1})$, per la seconda colonna di A. Otteniamo: $a_{12}\mathcal{S}_{11} + a_{22}\mathcal{S}_{21} + a_{32}\mathcal{S}_{31} + \cdots + a_{n2}\mathcal{S}_{n1}$, cioè lo sviluppo del determinante rispetto alla prima colonna di una matrice in cui le colonne dalla seconda alla $n - esima$ sono uguali, nell'ordine, a quelle di A e in cui la prima colonna è uguale alla seconda:

$$\begin{pmatrix} a_{12} & a_{12} & \ldots & a_{1n} \\ a_{22} & a_{22} & \ldots & a_{2n} \\ a_{32} & a_{32} & \ldots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n2} & \ldots & a_{nn} \end{pmatrix}.$$

La matrice in questione ha due colonne uguali perciò il suo determinante è uguale a zero. La medesima situazione si ripete per ogni entrata di posto i, k della matrice $\mathcal{S}^t A$ con $i \neq k$. Infatti il prodotto della riga $i - esima$ di $\mathcal{S}^t$ per la colonna $k-esima$ di A $(i \neq k)$ è il prodotto della colonna i-esima di $\mathcal{S}$ per la colonna $k-esima$ di A cioè: $a_{1k}\mathcal{S}_{1i} + a_{2k}\mathcal{S}_{2i} + a_{3k}\mathcal{S}_{3i} + \cdots + a_{nk}\mathcal{S}_{ni}$ e questo è il determinante di una matrice le cui colonne coincidono con quelle di A fuorché la $i-esima$ che è uguale alla $k-esima$: il determinante di una siffatta matrice (due colonne sono eguali) è nullo!

Osservazione 8.3.3 Data $A \in \mathcal{M}_n(\mathbb{R})$ chi è il determinante di A^t? In un caso la risposta balza agli occhi: se $\det(A) = 0$ questo significa che le colonne (o le righe) di A sono linearmente dipendenti, quindi le righe (o le colonne) di A^t sono pure linearmente dipendenti. Dunque $\det(A^t) = 0$.

D'altro canto se A è una matrice in $\mathcal{M}_1(\mathbb{R})$ essa coincide con la propria trasposta, dunque $\det(A) = \det(A^t)$. Un calcolo facile mostra che $\det(A) = \det(A^t)$ se $A \in \mathcal{M}_2(\mathbb{R})$. Sia ora $A \in \mathcal{M}_3(\mathbb{R})$ e sviluppiamo il determinante di A secondo gli elementi della prima riga. Sviluppando ora il determinante di A^t secondo gli elementi della prima colonna è facile vedere che si ottiene lo stesso determinante. Per lo stesso motivo l'identità $\det(A) = \det(A^t)$ è vera per ogni $A \in \mathcal{M}_n(\mathbb{R})$.

8.4 Cambiamenti di base

Proponiamo due semplici esempi/problemi per giustificare l'argomento di questo paragrafo:

Problema 1. In $\mathbb{R}^2$ consideriamo due basi: $\mathcal{V} = \{v_1 = (1,2),\ v_2 = (0,-1)\}$ e $\mathcal{W} = \{w_1 = (1,1),\ w_2 = (3,1)\}$. Poiché sono delle basi, ogni vettore di $\mathbb{R}^2$ si scrive in modo unico come combinazione lineare dei vettori di ciascuno dei due insiemi. Così, se prendiamo, ad esempio, $(-1,2) \in \mathbb{R}^2$, abbiamo l'unica scrittura $(-1,2) = (-1)(1,2) + (-4)(0,-1)$, quindi possiamo univocamente individuare $(-1,2)$ nella base v_1, v_2 mediante le coordinate $(-1,-4)_{\mathcal{V}}$. Analogamente, sempre lo stesso vettore $(-1,2)$ è espresso nella base $\{w_1, w_2\}$ dalle coordinate $(\frac{7}{2}, -\frac{3}{2})_{\mathcal{W}}$. Come possiamo passare da una scrittura all'altra? È possibile costruire una macchina che ci consenta di ottenere le coordinate di un vettore rispetto ad una data base conoscendo le sue coordinate rispetto ad una base diversa?

Ricordiamo che, fissata la base canonica $\mathcal{E}$: $\{e_1 = (1,0,\ldots,0),\ e_2 = (0,1,0,\ldots,0),\ \ldots, e_n = (0,0,\ldots,0,1)\}$ di $\mathbb{R}^n$, per ogni vettore v di $\mathbb{R}^n$ le coordinate di v rispetto alla base canonica coincidono con le entrate di v.

Problema 2. Dobbiamo descrivere un'applicazione lineare f tra due spazi vettoriali reali di dimensione finita V e W. Come abbiamo fatto finora? Abbiamo fissato una base del dominio ed una del codominio e abbiamo costruito una matrice tale che, scritto un vettore v di V nelle coordinate della base scelta del dominio, allora il prodotto della matrice per la matrice colonna data dalle coordinate di v, fornisce le coordinate del vettore immagine, $f(v)$, nella base scelta del codominio. Se ora cambiassimo base sia nel dominio che nel codominio, cosa succederebbe alla matrice associata all'applicazione f? Cambierebbe.

Quindi saper descrivere un'applicazione lineare f rispetto a basi diverse è molto importante in quanto non vi sono in generale basi privilegiate, ma potrebbero esserci basi rispetto alle quali la matrice associata all'applicazione lineare f risulti più semplice (come visto in 6.3.4).

In questo paragrafo vogliamo risolvere i problemi 1 e 2.

Sia V uno spazio vettoriale di dimensione n e consideriamo l'endomorfismo di V dato dalla applicazione lineare identica, cioè dall'applicazione lineare che ad ogni vettore $v \in V$ associa il vettore v stesso:

$$\begin{array}{rccl} id_V: & V & \longrightarrow & V \\ & v & \longmapsto & v = id_V(v). \end{array}$$

Sia $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ una base di V. Abbiamo visto nella Lezione 7 che la matrice associata a id_V rispetto alla base $\mathcal{V}$ (sia nel dominio che nel codominio) è la matrice identica $I_n \in \mathcal{M}_n(\mathbb{R})$.

Supponiamo ora di voler cambiare la base nel codominio, cioè la base del dominio sarà sempre $\mathcal{V}$, mentre nel codominio prenderemo un'altra base: $\mathcal{W} = \{w_1, w_2, \ldots, w_n\}$. Come sarà la matrice associata all'applicazione id_V rispetto alle nuove basi?

Per fare un esempio prendiamo $n = 3$ e supponiamo che i vettori della base $\mathcal{W}$ si scrivano, rispetto ai vettori della base $\mathcal{V}$, come segue: $w_1 = 2v_1 - v_3, w_2 = v_3, w_3 = v_1 - v_2$. Una cosa balza subito all'occhio: saremmo già in grado di scrivere la matrice associata all'applicazione identica rispetto alla base w_1, w_2, w_3 nel dominio e alla base v_1, v_2, v_3 nel codominio. In effetti la prima colonna di tale matrice è data dalle coordinate di $id_V(w_1) = w_1$ rispetto a v_1, v_2, v_3: $w_1 = 2v_1 - v_3$, cioè $id_V(w_1) = (2, 0, -1)_{\mathcal{V}}$. Lo stesso ragionamento vale per w_2 e w_3: la matrice della applicazione identica rispetto alla base $\mathcal{W}$ nel dominio e $\mathcal{V}$ nel codominio risulta allora essere

$$A_{(id_V, \mathcal{W}, \mathcal{V})} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Attenzione, però: il nostro problema di partenza non era questo! Tuttavia la matrice appena trovata (con poca fatica) ci sarà di aiuto. Vediamo perché. Osserviamo, intanto che $id_V \circ id_V = id_V$ e consideriamo il seguente diagramma:

$$(V, \mathcal{V}) \xrightarrow{id} (V, \mathcal{W}) \xrightarrow{id} (V, \mathcal{V})$$

che vogliamo tradurre in termini matriciali. Dovremmo quindi determinare $C = BA$, con $A, B, C \in \mathcal{M}_3(\mathbb{R})$, dove A è la matrice della applicazione identica rispetto alla base $\mathcal{V}$ nel dominio e $\mathcal{W}$ nel codominio, B è la matrice associata all'applicazione identica rispetto alla base $\mathcal{W}$ nel dominio e $\mathcal{V}$ nel codominio e quindi C è la matrice dell'applicazione identica rispetto alla base $\mathcal{V}$ nel dominio e nel codominio. Ora, per quanto osservato sopra, $C = I_3$, e $B = A_{(id_V, \mathcal{W}, \mathcal{V})}$. Pertanto la matrice A, cioè la matrice che stavamo cercando, è l'inversa di $B = A_{(id_V, \mathcal{W}, \mathcal{V})}$:

$$I_3 = BA,$$

quindi siamo in grado di calcolarla! Ma che cosa rappresenta la matrice A? È la matrice che trasforma un vettore scritto in coordinate rispetto alla base $\mathcal{V}$, nello stesso vettore scritto in coordinate rispetto alla base $\mathcal{W}$:

$$A = A_{(id_V, \mathcal{V}, \mathcal{W})} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Per questo la matrice A si chiama **matrice del cambiamento di base dalla base $\mathcal{V}$ alla base $\mathcal{W}$**.

Se a questo punto vogliamo conoscere le coordinate rispetto alla base $\mathcal{W}$ del vettore $v_1 + 2v_2 - 3v_3 = (1, 2, -3)_{\mathcal{V}}$ basta applicare la matrice del cambiamento di base (dalla base $\mathcal{V}$ alla base $\mathcal{W}$) al vettore $(1, 2, -3)$:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -\frac{3}{2} \\ -2 \end{pmatrix},$$

cioè $(1, 2, -3)_{\mathcal{V}} = (\frac{3}{2}, -\frac{3}{2}, -2)_{\mathcal{W}}$.

Conclusione del Problema 1. Possiamo dedurre un metodo per risolvere il problema 1. Abbiamo uno spazio vettoriale e due sue basi: $\mathcal{V}$ e $\mathcal{W}$. La matrice di passaggio dalla base $\mathcal{W}$ alla base $\mathcal{V}$, cioè la matrice $A \in \mathcal{M}_n(\mathbb{R})$

che ci permette di determinare le coordinate nella base $\mathcal{V}$ di un vettore espresso in coordinate rispetto alla base $\mathcal{W}$, è data dalla matrice nelle cui colonne si trovano, nell'ordine, le coordinate dei vettori $w_1, \ldots, w_n$ rispetto alla base $\mathcal{V}$. D'altro canto, se vogliamo la matrice del cambiamento di base dalla base $\mathcal{V}$ alla base $\mathcal{W}$, allora prenderemo A^{-1}.

Se abbiamo, ad esempio, in $\mathbb{R}^2$ la base $\mathcal{V}$ data da $v_1 = (1,1), v_2 = (1,-1)$ e la base $\mathcal{W}$ espressa nella base $\mathcal{V}$ come $w_1 = v_1 - 2v_2, w_2 = -v_1 + 3v_2$, allora la matrice di passaggio dalla base $\mathcal{W}$ alla base $\mathcal{V}$ è

$$A = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix},$$

mentre la sua inversa è la matrice di passaggio dalla base $\mathcal{V}$ alla base $\mathcal{W}$:

$$A^{-1} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}.$$

Nel Problema 1 la situazione è leggermente diversa poiché i vettori delle due basi date non sono espressi gli uni in coordinate rispetto agli altri. Si tratta di fare un passaggio in più e di determinare le coordinate dei vettori w_i nella base $\mathcal{V}$: $w_1 = (1,1) = \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1(1,2) + \alpha_2(0,-1) = (\alpha_1, 2\alpha_1 - \alpha_2)$ e quindi $\alpha_1 = 1$ e $\alpha_2 = 1$; $w_2 = (3,1) = \beta_1 v_1 + \beta_2 v_2 = \beta_1(1,2) + \beta_2(0,-1) = (\beta_1, 2\beta_1 - \beta_2)$ cioè $\beta_1 = 3$ e $\beta_2 = 5$. Quindi la matrice di passaggio dalla base $\mathcal{W}$ alla base $\mathcal{V}$ è

$$\begin{pmatrix} 1 & 3 \\ 1 & 5 \end{pmatrix}.$$

Non resta che prendere la sua inversa per calcolare la matrice di passaggio dalla base $\mathcal{V}$ alla base $\mathcal{W}$:

$$\begin{pmatrix} \frac{5}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Possiamo procedere anche in un modo diverso. In effetti sia i vettori della base $\mathcal{V}$ che i vettori della base $\mathcal{W}$ nel Problema 1 sono espressi in coordinate rispetto alla base canonica. Quindi i dati dell'esempio ci forniscono immediatamente la matrice di passaggio dalla base $\mathcal{V}$ alla base canonica $\mathbf{e}$: $T = \begin{pmatrix} 1 & 0 \\ 2 & -1 \end{pmatrix}$ e la matrice di passaggio dalla base $\mathbf{w}$ alla base $\mathbf{e}$: $S = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$. In termini di composizione possiamo allora scrivere la matrice di passaggio dalla base $\mathcal{V}$ alla base $\mathcal{W}$ come il prodotto $S^{-1}T$. Questo significa che stiamo componendo l'applicazione identica con se stessa, una volta descrivendola mediante T rispetto alla base $\mathcal{V}$ nel dominio ed $\mathbf{e}$ nel codominio, una volta mediante S^{-1} rispetto alla base $\mathbf{e}$ nel dominio e $\mathcal{W}$ nel codominio: il prodotto $S^{-1}T$ è dunque la matrice della applicazione identica rispetto alla base $\mathcal{V}$ nel dominio e $\mathcal{W}$ nel codominio, cioè quello che cercavamo:

$$S^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}; \qquad S^{-1}T = \begin{pmatrix} \frac{5}{2} & -\frac{3}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Passiamo ora al Problema 2. Sia $f \in \mathcal{L}in(V, W)$ e siano $\dim V = m$ e $\dim W = n$. Sia $F \in \mathcal{M}_{n,m}(\mathbb{R})$ la matrice associata a f rispetto ad una base fissata $\mathcal{V}$ di V e ad una base fissata $\mathcal{W}$ di W. Sia $H \in \mathcal{M}_m(\mathbb{R})$ la matrice del cambiamento di base dalla base $\mathcal{V}'$ alla base $\mathcal{V}$ di V e sia $K \in \mathcal{M}_n(\mathbb{R})$ la matrice del cambiamento di base dalla base $\mathcal{W}$ alla base $\mathcal{W}'$. Cerchiamo la matrice $F' \in \mathcal{M}_{n,m}(\mathbb{R})$ della applicazione f rispetto alla base $\mathcal{V}'$ nel dominio e alla base $\mathcal{W}'$ nel codominio. Moltiplicando F' per il vettore colonna delle coordinate di un vettore v di V nella base $\mathcal{V}'$ otterremo le coordinate di $f(v)$ nella base $\mathcal{W}'$ di W.

Partiamo dunque da un vettore di V espresso nella base $\mathcal{V}'$, tramite la matrice H lo trasformiamo nello stesso vettore espresso però in coordinate nella base $\mathcal{V}$, adesso possiamo applicare F che spedisce il vettore trovato nel vettore immagine tramite f, espresso in coordinate rispetto alla base $\mathcal{W}$ del codominio, a questo vettore applichiamo la matrice K che ci fornisce le sue coordinate nella base $\mathcal{W}'$ e a questo punto abbiamo finito. In termini di prodotto di matrici si ha:

$$F' = KFH.$$

In altre parole abbiamo scritto l'applicazione lineare f come composizione:

$$\begin{array}{ccc} (V, \mathcal{V}) & \xrightarrow{f} & (W, \mathcal{W}) \\ {\scriptstyle id_V} \uparrow & & \downarrow {\scriptstyle id_V} \\ (V, \mathcal{V}') & \xrightarrow{f} & (W, \mathcal{W}') \end{array}$$

Conclusione del Problema 2. Se la matrice $F \in \mathcal{M}_{n,m}(\mathbb{R})$ rappresenta una applicazione lineare $f \in \mathcal{L}in(V, W)$ tra gli spazi vettoriali V e W, di dimensione rispettivamente m e n, rispetto alle basi $\mathcal{V}$ e $\mathcal{W}$ e se F' è la matrice della stessa applicazione rispetto alle basi $\mathcal{V}'$ nel dominio e $\mathcal{W}'$ nel codominio, allora si ha:

$$F' = KFH$$

dove $H \in \mathcal{M}_m(\mathbb{R})$ è la matrice invertibile che rappresenta il cambiamento di base dalla base $\mathcal{V}'$ alla base $\mathcal{V}$, mentre $K \in \mathcal{M}_n(\mathbb{R})$ è la matrice invertibile che rappresenta il cambiamento di base dalla base $\mathcal{W}$ alla base $\mathcal{W}'$.

Esempio. Sia $\varphi \in \mathcal{L}in(\mathbb{R}^3, \mathbb{R}^2)$ l'applicazione lineare che rispetto alle basi canoniche in $\mathbb{R}^3$ e $\mathbb{R}^2$ è data dalla matrice

$$F = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \end{pmatrix}.$$

(Denoteremo con $\mathcal{E}$ sia la base canonica di $\mathbb{R}^3$ che la base canonica di $\mathbb{R}^2$). In particolare l'immagine mediante φ del vettore $(\alpha, \beta, \gamma) \in \mathbb{R}^3$, che nella base canonica ha coordinate (α, β, γ), è il vettore

$$F(\alpha\ \beta\ \gamma)^t = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \alpha - \beta \\ -\alpha + \beta + 2\gamma \end{pmatrix}.$$

Vogliamo ora scrivere la matrice associata a φ rispetto alla base di $\mathbb{R}^3$ data da $v_1 = (2, 1, 0), v_2 = (1, 1, 1), v_3 = (1, 0, 1)$ e alla base di $\mathbb{R}^2$ data da $w_1 = (1, 2), w_2 = (2, 1)$. La matrice del cambiamento di base dalla base $\mathcal{V}$ alla base canonica di $\mathbb{R}^3$ è la matrice

$$H = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

e la matrice del cambiamento di base da $\mathcal{W}$ ad $\mathcal{E}$ è

$$K^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

La matrice inversa è la matrice di passaggio dalla base canonica di $\mathbb{R}^2$ alla base $\mathcal{W}$:

$$K = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix}.$$

Così la matrice che cerchiamo è

$$F' = KFH = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & \frac{4}{3} & \frac{1}{3} \\ 1 & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}.$$

Esempio 8.4.1 Sia V uno spazio vettoriale di dimensione n e consideriamo lo spazio vettoriale degli endomorfismi $\mathcal{E}nd(V)$. Supponiamo che un endomorfismo φ ci venga dato nella base $\{v_1, v_2, \ldots, v_n\}$ dalla matrice $A \in \mathcal{M}_n(\mathbb{R})$ cioè scegliamo sia nel dominio che nel codominio la stessa base $\mathcal{V}$. Supponiamo di voler determinare la matrice della stessa applicazione rispetto ad una nuova base $\{w_1, w_2, \ldots, w_n\}$ di V (sia nel dominio che nel codominio). Se $H \in \mathcal{M}_n(\mathbb{R})$ è la matrice (invertibile) del cambiamento di base dalla base $\mathcal{W}$ alla base $\mathcal{V}$, allora la matrice che cerchiamo è $H^{-1}AH$.

Esempio 8.4.2 Data una base $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ di uno spazio vettoriale, qualsiasi matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ rappresenta la matrice di un cambiamento di base. Ma tra quali basi? E in che ordine? In effetti se consideriamo la matrice come la matrice di passaggio da una base $\mathcal{W}$ (che non conosciamo) alla base $\mathcal{V}$ allora le colonne di H rappresentano nell'ordine le coordinate nella base $\mathcal{V}$ dei vettori della base $\mathcal{W}$: $w_1, w_2, \ldots, w_n$.

D'altro canto possiamo pure interpretare H come la matrice del cambiamento di base dalla base $\mathcal{V}$ ad una base $\mathcal{W}$ che non conosciamo. Allora per determinare $\mathcal{W}$ basterà prendere la matrice inversa H^{-1} che rappresenterà la matrice del cambiamento di base dalla base $\mathcal{W}$ alla base conosciuta $\mathcal{V}$ e allora le sue colonne saranno le coordinate rispetto a $\mathcal{V}$ dei vettori (nell'ordine) della base $\mathcal{W}$.

In conclusione, fissata una base $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ di uno spazio vettoriale V, possiamo identificare le matrici invertibili di $\mathcal{M}_n(\mathbb{R})$ e l'insieme delle basi di V: ogni matrice invertibile H può essere vista come la matrice di un cambiamento di base dalla base $\mathcal{W}$ alla base $\mathcal{V}$, nel senso che le colonne di H sono nell'ordine le coordinate dei vettori $w_1, w_2, \ldots, w_n$.

Osservazione 8.4.3 Abbiamo già osservato che le operazioni elementari sulle righe di una matrice sono equivalenti a dei cambiamenti di base.

Infatti sia $L : V \to W$ un'applicazione lineare di matrice A rispetto ad una base $\mathcal{V}$ di V e ad una base $\mathcal{W}$ di W. Scambiare il vettore i-esimo con il vettore j-esimo della base del codominio W equivale a moltiplicare a sinistra la matrice A per la matrice $S_{i,j}$ ottenuta dalla matrice identica con l'operazione elementare $S_{i,j}$ che corrisponde a scambiare le righe corrispondenti della matrice A. Moltiplicare il vettore i-esimo della base del codominio per $\alpha \neq 0$ equivale a moltiplicare la matrice A a sinistra per la matrice $H_i(1/\alpha)$ (calcolare tale matrice per esercizio) che a sua volta corrisponde a moltiplicare la riga i-esima della matrice A per $\frac{1}{\alpha}$. Scegliere una nuova base del codominio in cui, ad esempio, l'i-esimo vettore è il vecchio i-esimo a cui si somma β volte il j-esimo e gli altri rimangono invariati corrisponde a sommare alla riga j-esima della matrice A la riga i-esima moltiplicata per $-\beta$, cioè a moltiplicare a sinistra per la matrice $H_{i,j}(-\beta)$ (determinarla per esercizio).

Analogamente scambiare l'i-esimo vettore della base del dominio V con il vettore j-esimo equivale a scambiare le colonne corrispondenti della matrice A. Moltiplicare il vettore i-esimo della base del dominio per $\alpha \neq 0$ equivale a moltiplicare la colonna i-esima per $\frac{1}{\alpha}$. Sommare al vettore i-esimo β volte il j-esimo e lasciare gli altri invariati corrisponde a sommare alla colonna j-esima della matrice A la colonna i-esima moltiplicata per $-\beta$.

Un cambiamento di base è un isomorfismo e il suo determinante è diverso da zero. Dunque le trasformazioni per riga (o per colonna se si cambia base nel dominio) sono trasformazioni invertibili. Presa una matrice quadrata A ed indicata con A' la matrice che si ottiene dalla matrice A mediante una trasformazione per riga (o colonna), si ha: $\det(A') = \alpha \det(A)$ dove $\alpha \neq 0$ è il determinante della matrice del cambiamento di base corrispondente alla trasformazione per riga (o per colonna) effettuata. Questo altera il determinante, ma non il fatto che sia o meno uguale a zero!

Definizione 8.4.4 *Due matrici $A, B \in \mathcal{M}_n(\mathbb{R})$ si dicono* **simili** *se rappresentano lo stesso endomorfismo rispetto a basi diverse, cioè, per quanto visto in 8.4.1, se esiste una matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ tale che $A = H^{-1}BH$.*

Nella definizione precedente le matrici A e B rappresentano lo stesso endomorfismo rispetto a basi diverse (la matrice A è 'legata' alla matrice B tramite un cambiamento di base dato dalla matrice H). Indichiamo l'insieme delle matrici invertibili di ordine n con $GL_n(\mathbb{R})$.

Osservazione 8.4.5 Segue immediatamente dalla definizione che matrici simili hanno lo stesso determinante. Infatti se A e B sono matrici quadrate di ordine n e sono simili, esiste una matrice invertibile $H \in \mathcal{M}_n(\mathbb{R})$ tale che $A = H^{-1}BH$. Di conseguenza $\det(A) = \det(H^{-1}BH) = \det(H^{-1})\det(B)\det(H)$ $= \frac{1}{\det(H)}\det(B)\det(H) = \det(B)$.

Osservazione 8.4.6 Il calcolo del determinante e la sua applicazione al calcolo del rango dei minori quadrati di una matrice forniscono un ulteriore metodo per il calcolo del rango di una matrice e quindi per la soluzione di un sistema lineare. Consideriamo il seguente esempio: in $\mathbb{R}^4$ si calcoli, al variare di $\alpha \in \mathbf{R}$, lo spazio delle soluzioni Σ_α del sistema lineare S_α nelle incognite x, y, z, w:

$$S_\alpha = \begin{cases} 2x + 3y - 4z - w = 0 \\ x + 2y - 3z = 0 \\ \alpha x + z - 2w = 0 \end{cases}.$$

S_α è un sistema omogeneo di 3 equazioni in 4 incognite. Scriviamo la matrice (completa=incompleta in questo caso) ad esso associata:

$$A_\alpha = \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ \alpha & 0 & 1 & -2 \end{pmatrix}$$

e consideriamo il minore individuato dalle prime tre colonne

$$\begin{pmatrix} 2 & 3 & -4 \\ 1 & 2 & -3 \\ \alpha & 0 & 1 \end{pmatrix}.$$

Il suo determinante è uguale a $-\alpha + 1$ quindi per ogni $\alpha \neq 1$ esso è diverso da zero! In questo caso la matrice A_α ha rango massimo (uguale a 3), quindi rappresenta una applicazione lineare di rango 3 di uno spazio di dimensione 4 in uno spazio di dimensione 3. Il sistema ammette soluzioni e lo spazio delle soluzioni è il nucleo dell'applicazione lineare individuata dalla matrice A_α e ha pertanto dimensione 1. Per determinare le soluzioni del sistema riduciamo la matrice A_α in forma a scalini per righe. Attraverso un certo numero di passaggi otteniamo la matrice:

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 1-\alpha & 2\alpha - 2 \end{pmatrix}$$

e, dunque, il sistema

$$\begin{cases} x + 2y - 3z = 0 \\ -y + 2z - w = 0 \\ (1-\alpha)z + (2\alpha - 2)w = 0. \end{cases}$$

Otteniamo $\Sigma_\alpha = \langle (0, 3, 2, 1) \rangle$.

Dobbiamo ora considerare il caso $\alpha = 1$. Attenzione: non stiamo dicendo che la matrice ha rango minore di 3 in questo caso! Vi sono altri minori quadrati di ordine 3 di cui non abbiamo calcolato il determinante. Può ancora succedere che la matrice abbia rango 3: per avere rango minore di 3, tutti i minori di ordine 3 debbono avere determinante uguale a 0. Il calcolo delle soluzioni del sistema S_1 si può fare direttamente per sostituzione. Abbiamo il sistema:

$$S_1 = \begin{cases} 2x + 3y - 4z - w = 0 \\ x + 2y - 3z = 0 \\ x + z - 2w = 0. \end{cases}$$

Calcoliamo allora il rango del sistema col metodo della riduzione per righe:

$$A_1 = \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ 1 & 0 & 1 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & -2 & 4 & -2 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & -2 & 4 & -2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La matrice A_1 ha dunque rango 2. Lo spazio delle sue soluzioni è uno spazio vettoriale di dimensione 2: le variabili libere sono z e w, dunque scegliendo $z = 1, w = 0$ si trova la soluzione $x = -1, y = 2, z = 1, w = 0$, mentre scegliendo $z = 0, w = 1$, si ha $x = 2, y = -1, z = 0, w = 1$. L'insieme delle soluzioni è: $\Sigma_1 = \langle(-1, 2, 1, 0), (2, -1, 0, 1)\rangle$.

8.5 Esercizi svolti

Esercizio 8.5.1 Sia $id : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione identica. Si consideri la base $\mathcal{B}$ di $\mathbb{R}^3$: $\mathcal{B} = \{w_1 = (1, 1, 0), w_2 = (1, -1, 0), w_3 = (0, 1, 1)\}$. Si scriva la matrice associata all'applicazione id fissando:

1) la base $\mathcal{B}$ nel dominio e la base canonica nel codominio;

2) la base canonica nel dominio e la base $\mathcal{B}$ nel codominio;

3) la base $\mathcal{B}$ nel dominio e nel codominio.

Svolgimento. L'applicazione id manda ogni vettore di $\mathbb{R}^3$ in se stesso: $id(v) = v$ per ogni $v \in \mathbb{R}^3$. Nel caso 1) abbiamo dunque $id(w_i) = w_i$ per ogni $i = 1, 2, 3$ e, poiché le coordinate di un vettore rispetto alla base canonica non sono che le sue componenti, la matrice associata all'applicazione id rispetto alla base $\mathcal{B}$ nel dominio e alla base canonica nel codominio è:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nel caso 2) abbiamo $id(e_i) = e_i$ per ogni vettore e_i della base canonica e dobbiamo esprimere i vettori e_i in coordinate rispetto alla base $\mathcal{B}$. Potremmo procedere direttamente determinando, per ogni $i = 1, 2, 3$, i numeri reali α, β, γ tali che $e_1 = \alpha w_1 + \beta w_2 + \gamma w_3$ e così per e_2 ed e_3. Tuttavia (astuti!) osserviamo che la matrice che ha sulle colonne le componenti dei vettori e_i rispetto ai vettori w_j è la matrice del cambiamento di base dalla base canonica alla base $\mathcal{B}$ pertanto essa è la matrice inversa della matrice determinata in 1):

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

Nel caso 3) abbiamo, banalmente, $id(w_1) = 1w_1 + 0w_2 + 0w_3$, $id(w_2) = 0w_1 + 1w_2 + 0w_3$, $id(w_3) = 0w_1 + 0w_2 + 1w_3$. Pertanto la matrice richiesta è la matrice I_3.

Esercizio 8.5.2 Sia $\mathcal{B} = \{v_1 = (1, 1), v_2 = (1, 0)\}$ una base di $\mathbb{R}^2$ e sia $\mathcal{E}$ la base canonica di $\mathbb{R}^3$. Sia $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ l'applicazione lineare associata, rispetto alla base $\mathcal{B}$ del dominio e alla base $\mathcal{E}$ del codominio, alla matrice

$$F = \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ 2 & 0 \end{pmatrix}.$$

Determinare la matrice associata all'applicazione lineare f rispetto alla base canonica $\mathcal{E}'$ di $\mathbb{R}^2$ e alla base $\mathcal{B}' = \{w_1 = (1, 1, 1), w_2 = (1, 1, 0), w_3 = (1, 0, 0)\}$ di $\mathbb{R}^3$.

Svolgimento. L'esercizio consiste nel tradurre in termini matriciali il seguente diagramma:

$$(\mathbb{R}^2, \mathcal{E}') \xrightarrow{id} (\mathbb{R}^2, \mathcal{B}) \xrightarrow{f} (\mathbb{R}^3, \mathcal{E}) \xrightarrow{id} (\mathbb{R}^3, \mathcal{B}').$$

La composizione di funzioni illustrata nel diagramma produce infatti il seguente effetto: si parte da un vettore di $\mathbb{R}^2$ espresso in coordinate rispetto alla base canonica di $\mathbb{R}^2$, lo si esprime in coordinate rispetto alla base $\mathcal{B}$, si applica ad esso f e si ottengono le coordinate del vettore immagine nella base canonica di $\mathbb{R}^3$; infine si determinano le coordinate del vettore immagine nella base $\mathcal{B}'$. La matrice della applicazione composta $id \circ f \circ id$ è la matrice richiesta dall'esercizio.

Indichiamo con $M_{\mathcal{E}'}^{\mathcal{B}}$ la matrice del cambiamento di base dalla base $\mathcal{E}'$ alla base $\mathcal{B}$, cioè la matrice associata all'applicazione identica di $\mathbb{R}^2$ rispetto alla base canonica del dominio e alla base $\mathcal{B}$ del codominio, e con $M_{\mathcal{E}}^{\mathcal{B}'}$ la matrice del cambiamento di base dalla base canonica alla base $\mathcal{B}'$ di $\mathbb{R}^3$. Allora $M_{\mathcal{E}'}^{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$. Analogamente $M_{\mathcal{E}}^{\mathcal{B}'} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$.

Così la matrice richiesta è:

$$M_{\mathcal{E}}^{\mathcal{B}'} F M_{\mathcal{E}'}^{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \\ 3 & -5 \end{pmatrix}.$$

Esercizio 8.5.3 Stabilire se i vettori $v_1 = (1, 3, 4)$, $v_2 = (4, 3, -2)$, $v_3 = (2, 3, 0)$ di $\mathbb{R}^3$ sono linearmente indipendenti.

Svolgimento. Abbiamo già risolto questo tipo di problema utilizzando la definizione di vettori linearmente indipendenti. Vogliamo proporre una soluzione diversa e decisamente più rapida che fa uso della nozione di determinante. Costruiamo la matrice che abbia come vettori riga i vettori v_1, v_2, v_3:

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 4 & 3 & -2 \\ 2 & 3 & 0 \end{pmatrix}.$$

A questo punto dire che i vettori v_1, v_2, v_3 sono linearmente indipendenti equivale a dire che le righe della matrice A sono linearmente indipendenti cioè che la matrice A è invertibile. Del resto la matrice A è invertibile se e solo se il suo determinante è non nullo. Dunque i vettori v_1, v_2, v_3 sono linearmente indipendenti se e solo se $\det(A) \neq 0$. Non ci resta che calcolare $\det(A)$: sviluppando il determinante di A secondo gli elementi della terza colonna otteniamo

$$\det(A) = 4(12 - 6) + 2(3 - 6) = 24 - 6 = 18.$$

Possiamo concludere che i vettori v_1, v_2 e v_3 sono linearmente indipendenti.

Esercizio 8.5.4 Calcolare il rango della matrice $M = \begin{pmatrix} 1 & 0 & -1 & 4 \\ -1 & 1 & 0 & 2 \\ -1 & 0 & 1 & 3 \end{pmatrix}$.

Svolgimento. Non vogliamo utilizzare il metodo della riduzione di una matrice in forma a scalini per righe ma la caratterizzazione del rango di una matrice come massimo ordine di minori invertibili della matrice. Il rango della matrice M è minore o uguale a 3. Sarà uguale a 3 se riusciremo a trovare un minore di ordine 3 di M con determinante non nullo.

Cominciamo col minore che si ottiene da M eliminando la quarta colonna: $\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. Esso ha determinante nullo e quindi non fa al caso nostro.

Consideriamo allora il minore di M che si ottiene eliminando la prima colonna: esso ha determinante uguale a 7. Dunque abbiamo individuato un minore invertibile di M di ordine 3. M ha pertanto rango 3.

Esercizio 8.5.5 In $\mathbb{R}^3$ siano dati i vettori $v_1 = (2, t, 1)$, $v_2 = (-1, 1, 0)$, $v_3 = (1, 1, t)$ dove t è un parametro reale. Si consideri l'endomorfismo $f_t : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definito da: $f_t(e_1) = v_1$, $f_t(e_2) = v_2$, $f_t(e_3) = v_3$, dove $\{e_1, e_2, e_3\}$ è la base canonica di $\mathbb{R}^3$. Esistono valori del parametro t per i quali l'applicazione f_t è invertibile? In caso affermativo, per uno di questi valori si determini la matrice che rappresenta l'applicazione inversa di f_t rispetto alla base canonica.

Svolgimento. Costruiamo la matrice associata all'applicazione lineare f_t rispetto alla base canonica di $\mathbb{R}^3$:

$$F_t = \begin{pmatrix} 2 & -1 & 1 \\ t & 1 & 1 \\ 1 & 0 & t \end{pmatrix}.$$

Sviluppiamo il determinante di F_t rispetto alla seconda colonna:

$\det(F_t) = 1(t^2 - 1) + (2t - 1) = t^2 + 2t - 2$. Allora per ogni $t \in \mathbb{R}$ tale che $t^2 + 2t - 2 \neq 0$, vale a dire per ogni $t \neq -1 \pm \sqrt{3}$, $\det(F_t) \neq 0$ e l'applicazione f_t è invertibile. Viceversa, se $t = -1 + \sqrt{3}$ oppure $t = -1 - \sqrt{3}$ l'applicazione f_t non è invertibile.

Poniamo ora $t = 1$ e consideriamo la matrice F_1 associata a f_1 rispetto alla base canonica. Si ha: $\det(F_1) = 1$ e

$$F_1^{-1} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$

8.6 Esercizi proposti

Esercizio 8.6.1 Calcolare il determinante della seguente matrice:

$$A = \begin{pmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 2 \end{pmatrix}.$$

Esercizio 8.6.2 Sia $A \in \mathcal{M}_{15}(\mathbb{R})$ una matrice tale che $A = -A^t$ (*matrice antisimmetrica*). Quanto vale $\det(A)$?

Esercizio 8.6.3 Costruire due matrici quadrate A e B tali che $\det(A) + \det(B) \neq 0$ e $\det(A + B) = 0$.

Esercizio 8.6.4 Scrivere la matrice del cambiamento di base dalla base $\mathcal{B} = \{(1, 2, 1), (1, 1, 1), (0, 1, -1)\}$ alla base $\mathcal{B}' = \{(-1, 1, 0), (1, -1, 1), (2, 1, 3)\}$ di $\mathbb{R}^3$.

Esercizio 8.6.5 Si consideri l'applicazione lineare $D : \mathbb{R}^3[x] \longrightarrow \mathbb{R}^3[x]$ che ad ogni polinomio a coefficienti reali di grado minore o uguale a 3 nella variabile x associa la sua derivata rispetto ad x. Scrivere la matrice associata a D rispetto alla base $\mathcal{C} = \{1, x, x^2, x^3\}$ nel dominio e alla base $\mathcal{B} = \{x^3, x^3 + x^2, x^3 + x^2 + x, x^3 + x^2 + x + 1\}$ nel codominio. Calcolare il determinante della matrice ottenuta.

Lezione 9

Matrici diagonalizzabili

Tratteremo ora di endomorfismi di uno spazio vettoriale V di dimensione n. Sappiamo che, scelta una base di V, ad ogni endomorfismo resta associata una matrice $A \in \mathcal{M}_n(\mathbb{R})$. D'altro canto sappiamo anche che un cambiamento di base trasforma la matrice dell'endomorfismo in una matrice simile $B \in \mathcal{M}_n(\mathbb{R})$ per cui $B = H^{-1}AH$ con $H \in GL_n(\mathbb{R})$ (8.4.4). Ci porremo due problemi diversi:

1) trovare una base rispetto alla quale la matrice del nostro endomorfismo abbia la forma più semplice possibile;

2) trovare un modo per decidere se due matrici siano o non siano simili (se, cioè, rappresentino o meno lo stesso endomorfismo).

Le risposte che daremo sono parziali, ma sufficienti per affrontare uno dei problemi descritti nell'introduzione. Per una risposta completa bisognerebbe ricorrere alla teoria di Jordan.

Sottolineiamo che useremo in maniera equivalente matrici quadrate ed endomorfismi.

9.1 Autovalori e autovettori

Consideriamo $\varphi \in \mathcal{E}nd(V)$.

Definizione 9.1.1 *Un numero reale λ si dice un autovalore dell'endomorfismo φ se esiste un vettore $v \neq \mathbf{0}_V$ tale che $\varphi(v) = \lambda v$. Il vettore v si dice allora autovettore di φ di autovalore λ.*

Si noti che abbiamo richiesto che un autovettore sia diverso dal vettore nullo: in effetti se accettassimo anche $\mathbf{0}_V$ come possibile autovettore allora si avrebbe $\mathbf{0}_V = \varphi(\mathbf{0}_V) = \alpha\mathbf{0}_V$ per ogni reale α. Dunque ogni numero reale sarebbe un autovalore e la definizione non avrebbe tanto senso. Si osservi inoltre che un vettore non nullo di V è un autovettore di φ di autovalore zero se e solo se appartiene al nucleo di φ.

Sia ora $\lambda \in \mathbb{R}$ un autovalore dell'endomorfismo $\varphi \in \mathcal{E}nd(V)$. Consideriamo l'insieme V_λ^* degli autovettori di autovalore λ (non è vuoto):

$$V_\lambda^* = \{v \in V \mid v \neq \mathbf{0}_V \, , \, \varphi(v) = \lambda v\}.$$

Aggiungiamo a tale insieme (che non può essere uno spazio vettoriale perché non contiene il vettore nullo) il vettore nullo: $V_\lambda = V_\lambda^* \cup \{\mathbf{0}_V\} = \{v \in V \mid \varphi(v) = \lambda v\}$. Allora

Proposizione 9.1.2 *L'insieme V_λ è un sottospazio vettoriale di V.*

Dimostrazione. Siano v_1 e v_2 due elementi di V_λ (i.e. $\varphi(v_1) = \lambda v_1$, $\varphi(v_2) = \lambda v_2$), dobbiamo vedere che pure $v_1 + v_2$ appartiene a V_λ. Calcoliamo quindi $\varphi(v_1 + v_2)$ che, per la linearità di φ, è uguale a $\varphi(v_1) + \varphi(v_2) = \lambda v_1 + \lambda v_2 = \lambda(v_1 + v_2)$. Dunque $v_1 + v_2$ è un autovettore di autovalore λ oppure il vettore nullo, in ogni caso $v_1 + v_2 \in V_\lambda$. Siano ora $\alpha \in \mathbb{R}$ e $v \in V_\lambda$, allora $\varphi(\alpha v) = \alpha\varphi(v) = \alpha(\lambda v)$, essendo φ lineare e $v \in V_\lambda$; del resto $\alpha(\lambda v) = \lambda(\alpha v)$ per le proprietà del prodotto per scalari, così $\alpha v \in V_\lambda$. Dunque V_λ è un sottospazio vettoriale di V. **C. V. D.**

Definizione 9.1.3 *Dati $\varphi \in \mathcal{E}nd(V)$ ed un suo autovalore λ, il sottospazio vettoriale V_λ è detto l'autospazio di φ associato all'autovalore λ.*

Per ora tutto è formalmente chiaro, ma finché non sappiamo calcolare gli autovalori di un endomorfismo non possiamo fare nulla di concreto. Innanzitutto osserviamo che, scelta una base $\mathcal{V} = \{v_1, \ldots, v_n\}$ di V ($\dim V = n$) e indicata con $A \in \mathcal{M}_n(\mathbb{R})$ la matrice associata a φ rispetto a questa base sia nel dominio che nel codominio di φ, dire che un vettore v è un autovettore di autovalore λ, significa che per la $n-upla$ delle sue coordinate $(a_1, \ldots, a_n)$ nella base $\mathcal{V}$ si ha

$$\lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} = A \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Questa condizione è indipendente dalla base scelta nel senso che, se $H \in GL_n(\mathbb{R})$ è la matrice del cambiamento di base dalla base $\mathcal{V}$ ad una base $\mathcal{W}$, allora le coordinate $H\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ sono le coordinate di v nella base $\mathcal{W}$ e soddisfano l'analoga relazione dove al posto di A troveremo la matrice B associata all'endomorfismo φ rispetto alla base $\mathcal{W}$. Cerchiamo di essere più precisi: per costruzione B è simile ad A attraverso la relazione esplicita: $B = HAH^{-1}$. Ne consegue che

$$B(H\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) = BH\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = (HAH^{-1}H)\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = HA\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = H\begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} =$$

$= \lambda(H\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix})$ cioè $H\begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}$ sono le coordinate nella base $\mathcal{W}$ di un autovettore di φ relativo all'autovalore λ.

Sia ora $\varphi \in \mathcal{E}nd(V)$ e sia $\mathcal{V}$ una base di V, $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$. Allora, se $A \in \mathcal{M}_n(\mathbb{R})$ è la matrice associata a φ rispetto alla base $\mathcal{V}$ (sia nel dominio che nel codominio di φ), possiamo considerare la matrice $A - tI_n$ (cioè la matrice che si ottiene sottraendo agli elementi diagonali di A l'indeterminata t). Il determinante di questa matrice è un polinomio nella variabile t:

$$P(t) = \det(A - tI_n)$$

detto **polinomio caratteristico di** φ. Tale polinomio sembrerebbe dipendere dalla matrice associata all'endomorfismo (e quindi dalla base scelta).

(Osserviamo tra parentesi che, ad essere precisi, la matrice $A - tI_n$ non è a coefficienti reali, ma a coefficienti nell'anello dei polinomi $\mathbb{R}[t]$ (che non è un campo) quindi non abbiamo definito cosa sia il determinante di una siffatta matrice. Tuttavia si può definire tale determinante in modo analogo a quanto fatto per le matrici a coefficienti in $\mathbb{R}$. In questo caso per matrici $p(t) \in \mathcal{M}_{1,1}(\mathbb{R}[t]) \cong \mathbb{R}[t]$ si avrà $\det p(t) = p(t)$ e per matrici quadrate di ordine superiore si utilizzerà sempre lo sviluppo per righe o per colonne.)

Osservazione 9.1.4 *Il polinomio caratteristico di φ non dipende dalla base scelta per costruire la matrice A.*

Dimostrazione. Per mostrare che il polinomio caratteristico di φ è indipendente dalla base scelta basta dimostrare che due matrici simili hanno lo stesso polinomio caratteristico. Fissata, infatti, una base di V diversa dalla precedente, la matrice associata a φ rispetto alla nuova base è una matrice simile alla matrice A. Siano dunque A, B due matrici simili: $B = H^{-1}AH$, con $H \in GL_n(\mathbb{R})$ matrice invertibile. Il polinomio caratteristico della matrice B è: $\det(B - tI_n) = \det(H^{-1}AH - tI_n)$. Osserviamo che tI_n è una matrice i cui soli elementi non nulli si trovano sulla diagonale e sono tutti uguali a t. Si vede facilmente che per ogni matrice quadrata T di ordine n, $(tI_n)T = t(I_nT) = tT = T(tI_n)$. Si ha pertanto: $\det(B - tI_n) = \det(H^{-1}AH - H^{-1}(tI_n)H) = \det(H^{-1}(A - tI_n)H) = \det H^{-1} \det(A - tI_n) \det H = \det(A - tI_n)$. **C.V.D.**

Per analogia parleremo dunque di polinomio caratteristico di una matrice $A \in \mathcal{M}_n(\mathbb{R})$,

$$\det(A - tI_n) = P_A(t).$$

Abbiamo visto che tale polinomio è lo stesso per ogni matrice simile ad A.

Proposizione 9.1.5 *Un numero reale λ è un autovalore dell'endomorfismo φ se e solo se è radice del suo polinomio caratteristico.*

Dimostrazione. Dire che $\alpha \in \mathbb{R}$ è un autovalore di φ equivale a dire, in coordinate rispetto ad una base $\mathcal{V}$, che esiste un vettore $(t_1, \ldots, t_n) \neq \mathbf{0}_{\mathbb{R}^n}$ tale che

$$A \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \alpha \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix}$$

cioè

$$A \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \alpha I_n \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} \iff (A - \alpha I_n) \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

i.e., $\mathbf{0}_{\mathbb{R}^n} \neq (t_1, \ldots, t_n) \in \operatorname{Ker}(A - \alpha I_n)$. Osserviamo che la matrice quadrata $(A - \alpha I_n)$ ha nucleo non banale se e solo se il suo rango non è massimo (se tale rango fosse massimo $\mathrm{Ker}(A-\alpha I_n)$ conterrebbe solo il vettore nullo). In altre parole α è autovalore di φ se e solo se $A-\alpha I_n$ non è invertibile, se e solo se, quindi, il suo determinante è uguale a zero: $\det(A-\alpha I_n)=0$. Ma tale determinante è proprio il polinomio caratteristico di A, $P_A(t)$, calcolato in α: $P_A(\alpha)$. Cosicché α è un autovalore di φ se e solo se $P_A(\alpha)=0$. **C.V.D.**

Osservazione 9.1.6 i) Esistono naturalmente matrici che non hanno autovalori reali! Consideriamo, ad esempio, la matrice $A=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è il polinomio

$$\det(A-\lambda I_2)=\det\begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix}=\lambda^2+1$$

che non ha radici reali.

ii) Il polinomio caratteristico di un endomorfismo φ di uno spazio vettoriale di dimensione n ha grado n, vale a dire, indicata con A la matrice associata a φ rispetto ad una base fissata, il polinomio $\det(A-\lambda I_n)$ è un polinomio di grado n nella variabile λ. Questo significa che il grado massimo con cui compare la variabile λ nello sviluppo del determinante della matrice $A-\lambda I_n$ è n. Per renderci conto di questo calcoliamo $\det(A-\lambda I_n)$ sviluppando il determinante rispetto alla prima riga: il termine di grado massimo (in λ) si ottiene moltiplicando il maggior numero possibile di termini contenenti λ. Dal momento che λ compare, con grado 1 e coefficiente -1, solo sulla diagonale della matrice e che gli elementi diagonali sono n, si deduce facilmente che il grado massimo con cui λ compare nella espressione del determinante è n e che il coefficiente di λ^n è esattamente $(-1)^n$.

Possiamo dire qualcosa sul termine noto del polinomio caratteristico? In questo caso dobbiamo considerare soltanto i termini che non contengono λ: basterebbe calcolare il polinomio caratteristico e porre $\lambda=0$ o, equivalentemente, porre $\lambda=0$ e calcolare il polinomio caratteristico. Ma quando $\lambda=0$ si ha $\det(A-0I_n)=\det(A)$. Dunque il termine noto del polinomio caratteristico della matrice A è $\det(A)$. Poiché matrici simili hanno lo stesso polinomio caratteristico, quanto appena osservato è coerente col fatto, già ottenuto in (come conseguenza diretta della definizione di similitudine di matrici) che matrici simili hanno lo stesso determinante.

Vediamo il significato di altri coefficienti del polinomio caratteristico. Consideriamo il coefficiente di λ^{n-1}: sviluppiamo, come prima, il determinante della matrice $A-\lambda I_n$ rispetto alla prima riga. Per chiarezza di notazioni poniamo $B = A - \lambda I_n$, allora:

$$\begin{aligned} \det(A - \lambda I_n) = \quad & (a_{11} - \lambda)\det(\mathcal{A}_{11} - \lambda I_{n-1}) - a_{12}\det(\mathcal{B}_{12}) + \\ & + a_{13}\det(\mathcal{B}_{13}) + \cdots + (-1)^{1+n} a_{1n}\det(\mathcal{B}_{1n}). \end{aligned} \tag{9.1}$$

Ricordiamo che le sottomatrici $\mathcal{B}_{1k}$ della matrice B con $k \neq 1$ si ottengono eliminando la prima riga e la colonna k-esima della matrice $B = A - \lambda I_n$, quindi eliminando due entrate della matrice $A - \lambda I_n$ contenenti la variabile λ: l'elemento di posto $1, 1$ e l'elemento di posto k, k. Inoltre le entrate a_{1k} della matrice A per $k \neq 1$ non contengono la variabile λ. Di conseguenza nella espressione (9.1) gli eventuali termini di grado $n-1$ nella variabile λ sono contenuti nell'addendo $(a_{11} - \lambda)\det(\mathcal{A}_{11} - \lambda I_{n-1})$. Proviamo allora che il coefficiente del termine di grado $n-1$ è $(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$. Per calcolarlo senza troppa fatica si può procedere per induzione su n: se $n = 1$, $A = (a_{11})$ e si vede immediatamente che il polinomio caratteristico di A è $a_{11} - \lambda$ e il termine di grado 0 è proprio $(-1)^0 a_{11} = a_{11}$. Dopodiché, assumiamo di conoscere il risultato per qualunque matrice di ordine minore di n e calcoliamo il coefficiente del termine di grado $n-1$ della nostra matrice B. Abbiamo detto che si tratta di calcolare i termini di grado $n-1$ in λ che compaiono nel prodotto $(a_{11} - \lambda)\det(\mathcal{A}_{11} - \lambda I_{n-1})$. Ora in $(a_{11} - \lambda)\det(\mathcal{A}_{11} - \lambda I_{n-1})$ i termini di grado $n-1$ sono quelli che si ottengono moltiplicando a_{11} per gli elementi di grado $n-1$ in $\det(\mathcal{A}_{11} - \lambda I_{n-1})$ e quelli che si ottengono moltiplicando $-\lambda$ per i termini di grado $n-2$ in $\det(\mathcal{A}_{11} - \lambda I_{n-1})$. Osserviamo che $\det(\mathcal{A}_{11} - \lambda I_{n-1})$ è precisamente il polinomio caratteristico della matrice $\mathcal{A}_{11}$ che è una matrice quadrata di ordine $n-1$. Quindi, per ipotesi induttiva e per quanto già osservato, i termini di grado $n-1$ nel polinomio caratteristico di A sono: $a_{11}((-1)^{n-1}\lambda^{n-1}) - \lambda((-1)^{n-2}(a_{22} + \cdots + a_{nn})\lambda^{n-2}) = (-1)^{n-1}\lambda^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$.

La somma $a_{11} + a_{22} + \cdots + a_{nn}$ è la somma degli elementi sulla diagonale di A. Tale somma è detta la **traccia** di A e si indica di solito con $\mathrm{tr}(A)$. Abbiamo quindi mostrato che il coefficiente del monomio di grado $n-1$ del polinomio caratteristico di una matrice quadrata A di ordine n è, a meno del segno (che dipende solo da n), la traccia della matrice A. In particolare matrici simili hanno la stessa traccia! (Esse hanno infatti lo stesso polinomio caratteristico.)

iii) Il polinomio caratteristico di un endomorfismo φ di uno spazio vettoriale n-dimensionale ha grado n, ma le sue radici possono anche non essere tutte distinte. Ne consegue che gli autovalori distinti di φ sono in numero minore o uguale a n.

Che relazione c'è tra autospazi relativi ad autovalori diversi? Si intersecano? La prossima proposizione risponde a questa domanda:

Proposizione 9.1.7 *Sia $\varphi \in \mathcal{E}nd(V)$ e sia $\mathcal{V}$ una base di V,* $\dim V = n$. *Siano $\alpha_1, \alpha_2, \dots, \alpha_k$ autovalori distinti di φ. Allora i relativi autospazi sono in somma diretta: $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \cdots \oplus V_{\alpha_k}$.*

Dimostrazione. Generalizzando quanto visto per le somme dirette di due sottospazi, basterà verificare che l'intersezione fra una somma qualsiasi di alcuni autospazi e qualunque autospazio che non compaia fra gli addendi di questa somma è banale. Procediamo per induzione su k. Innanzitutto mostriamo che la somma di due autospazi distinti è diretta: sia $v \in V_{\alpha_1} \cap V_{\alpha_2}$. Allora $\varphi(v) = \alpha_1 v$ e $\varphi(v) = \alpha_2 v$, i.e., $(\alpha_1 - \alpha_2)v = \mathbf{0}_V$ il che implica, essendo $\alpha_1 \neq \alpha_2$, $v = \mathbf{0}_V$.

Ora supponiamo che la somma dei j autospazi $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \cdots \oplus V_{\alpha_j}$ sia diretta e consideriamo V_{α_l} con $l \notin \{1, 2, \dots, j\}$. Sia poi $v \in (V_{\alpha_1} \oplus V_{\alpha_2} \oplus V_{\alpha_j}) \cap V_{\alpha_l}$. Si ha dunque: $v = v_1 + v_2 + \cdots + v_j = v_l$, con $v_s \in V_{\alpha_s}$, pertanto $\varphi(v_1 + v_2 + \cdots + v_j) = \varphi(v_l)$, ma φ è lineare e i vettori v_i sono autovettori, dunque

$$\varphi(v_1 + \cdots + v_j) = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_j v_j = \alpha_l v_l = \alpha_l v = \alpha_l v_1 + \cdots + \alpha_l v_j$$

cioè

$$\alpha_l v_1 + \alpha_l v_2 + \cdots + \alpha_l v_j = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_j v_j.$$

Dunque $(\alpha_l - \alpha_1)v_1 + (\alpha_l - \alpha_2)v_2 + \cdots + (\alpha_l - \alpha_j)v_j = \mathbf{0}_V$. Dal momento che, per ipotesi, la somma di $V_1, \dots, V_j$ è diretta, i vettori $v_1, \dots, v_j$ sono linearmente indipendenti, ma essendo $\alpha_l - \alpha_s \neq 0$ se $s \neq l$, si ha che tutti i vettori sono nulli: $v_1 = v_2 = \cdots = v_j = v_l = \mathbf{0}_V$. Quindi $(V_1 \oplus \cdots \oplus V_j) \cap V_l = 0_V$. **C. V. D.**

La Proposizione 9.1.7 ci assicura che autovettori relativi ad autovalori diversi sono linearmente indipendenti: essi appartengono infatti a spazi vettoriali in somma diretta.

9.2 Matrici/endomorfismi diagonalizzabili

Mettiamoci ora in una situazione "idilliaca": sia $\varphi \in \mathcal{E}nd(V)$ e supponiamo che la somma (diretta) degli autospazi $V_{\alpha_1} \oplus V_{\alpha_2} \oplus \cdots \oplus V_{\alpha_k}$ sia uguale a V, con $\alpha_1, \alpha_2, \dots, \alpha_k$ autovalori diversi. Allora se scegliamo una base di ciascun autospazio: $v_1^1, v_2^1, \dots, v_{n_1}^1$ per V_{α_1}, $v_1^2, v_2^2, \dots, v_{n_2}^2$ per V_{α_2} e così via, fino a $v_1^k, v_2^k, \dots, v_{n_k}^k$ per V_{α_k}, otteniamo una base di V come unione delle basi dei V_{α_i}. Adesso cerchiamo di scrivere la matrice dell'applicazione lineare φ rispetto a questa base. Cominciamo da v_1^1: v_1^1 appartiene al sottospazio V_{α_1}, quindi $\varphi(v_1^1) = \alpha_1 v_1^1$ che in coordinate nella base fissata si scrive come $\alpha_1 v_1^1 + 0v_2^1 + \cdots + 0v_{n_1}^1 + 0v_1^2 + \cdots + 0v_1^k + \cdots + 0v_{n_k}^k$. Per v_2^1 si ha $\varphi(v_2^1) = \alpha_1 v_2^1 = 0v_1^1 + \alpha_1 v_2^1 + 0v_3^1 + \cdots + 0v_{n_1}^1 + 0v_1^2 + \cdots + 0v_1^k + \cdots + 0v_{n_k}^k$. In definitiva la matrice associata a φ rispetto ad una base di autovettori è diagonale e gli elementi sulla diagonale sono gli autovalori:

$$D = \begin{pmatrix} \alpha_1 & 0 & 0 & \dots & 0 \\ 0 & \alpha_1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \alpha_k & 0 \\ 0 & 0 & \dots & 0 & \alpha_k \end{pmatrix}.$$

Poiché la matrice trovata è la matrice dell'endomorfismo rispetto ad una base di autovettori di φ, il suo polinomio caratteristico è quello dell'endomorfismo φ pertanto gli zeri del polinomio caratteristico di D debbono avere la stessa molteplicità degli zeri del polinomio caratteristico di φ; in altre parole α_i compare sulla diagonale di D un numero di volte pari alla sua molteplicità come radice del polinomio caratteristico di φ.

Definizione 9.2.1 *Sia λ un autovalore di un endomorfismo φ di uno spazio vettoriale V. Si dice molteplicità algebrica di λ la sua molteplicità come radice del polinomio caratteristico. Si dice molteplicità geometrica di λ la dimensione dell'autospazio relativo all'autovalore λ.*

Definizione 9.2.2 *Un endomorfismo $\varphi \in \mathcal{E}nd(V)$ si dice diagonalizzabile se V ammette una base di autovettori di φ. Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ si dice diagonalizzabile se è simile ad una matrice diagonale $D \in \mathcal{M}_n(\mathbb{R})$.*

Abbiamo appena visto che se la somma (diretta) degli autospazi coincide con lo spazio vettoriale V allora l'endomorfismo è diagonalizzabile e la matrice ad esso associata rispetto ad una di autovettori è diagonale.

Molteplicità algebrica e geometrica di un autovalore sono legate dal seguente risultato:

Teorema 9.2.3 *La molteplicità algebrica di un autovalore è sempre maggiore della sua molteplicità geometrica o uguale ad essa.*

Dimostrazione. Supponiamo che α sia un autovalore di molteplicità algebrica k. Allora il polinomio caratteristico si può scrivere come: $\det(A-zI_n) = (\alpha - z)^k Q(z)$, dove $Q(z)$ ha grado $n-k$ e $Q(\alpha) \neq 0$. Supponiamo ora che la molteplicità geometrica di α sia h cioè che l'autospazio relativo all'autovalore α abbia dimensione h e supponiamo $h > k$. Allora possiamo scegliere una base $v_1, v_2, \dots, v_k, v_{k+1}, \dots, v_h$ di V_α e completare questa base in una base $\mathcal{V}$ di V. La matrice associata a φ rispetto alla base $\mathcal{V}$ sarà del tipo

$$A' = \begin{pmatrix} \alpha & 0 & \dots & 0 & & \\ 0 & \alpha & \dots & \vdots & & \\ \vdots & 0 & \ddots & 0 & * & \\ & & & \alpha & & \\ \vdots & \vdots & 0 & 0 & & \\ \vdots & \vdots & \dots & \vdots & \text{B} & \\ 0 & 0 & \dots & 0 & & \end{pmatrix},$$

dove gli α sulla diagonale sono h e B è una matrice quadrata di ordine $n-h$. Il polinomio caratteristico della matrice A' è $(\alpha - z)^h \det(B - zI_{n-h})$. Ma A e A' sono matrici associate alla stessa applicazione lineare rispetto a basi diverse, pertanto sono simili e hanno lo stesso polinomio caratteristico. Questo non è possibile dal momento che α è radice del polinomio $(\alpha-z)^k Q(z)$ con molteplicità k ed è radice del polinomio $(\alpha - z)^h \det(B - zI_{n-h})$ con molteplicità almeno h, ma per ipotesi $h > k$. Assurdo. **CVD**

Teorema 9.2.4 *Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ (o, equivalentemente, un endomorfismo φ di uno spazio vettoriale n-dimensionale V) è diagonalizzabile (su $\mathbb{R}$) se e solo se valgono le seguenti due condizioni:*

1. *il polinomio caratteristico $p(z)$ di A (o di φ) si fattorizza in $\mathbb{R}[z]$ nel prodotto di polinomi di primo grado (non necessariamente distinti);*

2. *la molteplicità algebrica di ogni autovalore λ coincide con la sua molteplicità geometrica.*

Dimostrazione. $\Leftarrow$) Dire che $p(z)$ si fattorizza in $\mathbb{R}[z]$ nel prodotto di fattori lineari equivale a dire che $p(z)$ ha n radici reali (non necessariamente distinte) cioè che, se indichiamo con $\alpha_1, \dots, \alpha_k$ tali radici e con n_i la molteplicità di α_i, per $i = 1, \dots, k$, allora $n_1 + n_2 + \cdots + n_k = n$. Inoltre sappiamo che la somma $V_{\alpha_1} \oplus \cdots \oplus V_{\alpha_k}$ degli autospazi è diretta e, dal momento che per ogni

$i = 1, \dots, k$ la dimensione dell'autospazio V_{α_i} coincide con la molteplicità n_i dell'autovalore α_i, si ha $\dim(V_{\alpha_1} \oplus \cdots \oplus V_{\alpha_k}) = \dim(V_{\alpha_1}) + \ldots + \dim(V_{\alpha_k}) = n_1 + \cdots + n_k = n$ quindi $V_{\alpha_1} \oplus \cdots \oplus V_{\alpha_k} = V$.

$\Rightarrow$) Vogliamo ora mostrare che, viceversa, se un endomorfismo φ di uno spazio vettoriale V n-dimensionale è diagonalizzabile allora valgono le condizioni 1. e 2. Sia A la matrice associata a φ rispetto ad una base fissata di V. Se A è diagonalizzabile, per definizione essa è simile ad una matrice diagonale, cioè esiste una matrice invertibile H tale che $H^{-1}AH = D$ è diagonale. In particolare le matrici A e D hanno lo stesso polinomio caratteristico e, essendo D diagonale, tale polinomio caratteristico è il prodotto degli elementi diagonali della matrice $D - zI_n$, quindi esso è il prodotto di fattori lineari in z:

$$(\alpha_1 - z)^{n_1}(\alpha_2 - z)^{n_2} \cdots (\alpha_k - z)^{n_k}$$

dove gli α_i sono gli autovalori di D (e quindi di A) e $n_1 + n_2 + \cdots + n_k = n$, essendo n_i il numero di volte in cui α_i appare nella matrice diagonale D. Abbiamo così mostrato che il polinomio caratteristico $p(z)$ di una matrice diagonalizzabile si fattorizza in $\mathbb{R}[z]$ in polinomi di grado 1 (non necessariamente distinti). D'altra parte la matrice D è la matrice associata a φ rispetto ad una base $\mathcal{B}$ di autovettori di φ e la molteplicità algebrica dell'autovalore α_i, cioè il numero di volte in cui α_i compare sulla diagonale della matrice D, è uguale al numero di autovettori relativi all'autovalore α_i che appaiono nella base $\mathcal{B}$. Quindi la molteplicità geometrica di α è uguale alla sua molteplicità algebrica. **CVD**

Osservazione 9.2.5 *i*) Abbiamo già osservato che esistono matrici ad entrate reali prive di autovalori reali. Nello stesso modo esistono matrici il cui polinomio caratteristico, pur avendo alcune radici reali, non si fattorizza completamente in polinomi di grado 1 in $\mathbb{R}[x]$, cioè le sue radici non sono tutte reali. Consideriamo, ad esempio, la matrice

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Il suo polinomio caratteristico è $(3 - z)(z^2 + 1)$ che ha una sola radice reale $z = 3$ (e due radici complesse immaginarie).

ii) Anche se il polinomio caratteristico di una matrice è completamente fattorizzabile in $\mathbb{R}[x]$, può accadere che la matrice non sia diagonalizzabile.

Ad esempio prendiamo la matrice $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è $z^2 = (-z+0)^2$. Quindi B ha un solo autovalore: $\alpha = 0$ di molteplicità algebrica 2. Se B fosse diagonalizzabile sarebbe simile alla matrice diagonale con l'autovalore 0 sulla diagonale, cioè alla matrice nulla. Ma qualunque matrice simile alla matrice nulla è nulla: $H^{-1}\mathbf{O}_n H = \mathbf{O}_n$! Quindi la matrice B, che è non nulla, non può essere simile ad una diagonale, cioè B non è diagonalizzabile.

iii) Se A è una matrice diagonalizzabile la sua forma diagonale D è completamente determinata dai suoi autovalori: D ha come elementi (diagonali) gli autovalori di A in numero pari alla loro molteplicità algebrica. Ovviamente tale forma diagonale non è unica, è unica a meno di permutazioni degli elementi sulla diagonale.

iv) Se α è radice del polinomio caratteristico di un endomorfismo φ, allora φ ammette almeno un autovettore di autovalore α, quindi la dimensione dell'autospazio V_α è maggiore o uguale ad 1. In particolare se il polinomio caratteristico di una matrice è completamente fattorizzabile in $\mathbb{R}[x]$ e tutte le sue radici sono distinte, cioè se la molteplicità algebrica di ogni autovalore è uguale ad uno, allora la dimensione di ogni autospazio è esattamente uguale ad 1, dovendo essere maggiore o uguale ad 1 per quanto appena detto e minore o uguale ad 1 per quanto osservato sopra. Quindi in questo caso la molteplicità di ogni autovalore coincide con la dimensione dell'autospazio corrispondente. In questo caso, dunque, l'endomorfismo è diagonalizzabile.

Veniamo ora al metodo per verificare se un endomorfismo (o, equivalentemente, una matrice) sia o meno diagonalizzabile e trovare, in caso affermativo, una base di autovettori che lo diagonalizzi. Siano dati l'endomorfismo $\varphi \in \mathcal{E}nd(V)$ e la matrice $A \in \mathcal{M}_n(\mathbb{R})$ associata a φ rispetto ad una base fissata di V.

1) Si calcola il polinomio caratteristico $\det(A - zI_n)$. Se non è prodotto di fattori lineari in $\mathbb{R}[x]$, cioè se le sue radici non sono tutte reali, allora φ non è diagonalizzabile. Altrimenti: $\det(A-zI_n) = (\alpha_1-z)^{n_1}(\alpha_2-z)^{n_2}\cdots(\alpha_k-z)^{n_k}$, $\alpha_i \in \mathbb{R}$.

2) Si considerano gli autospazi V_{α_i}. Si ha $V_{\alpha_i} = Ker(A - \alpha_i I_n)$ quindi V_{α_i} è un sottospazio di V di dimensione $n - \mathrm{rg}(A - \alpha_i I_n)$, dato dalle soluzioni del sistema

$$(A - \alpha_i I_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \mathbf{0}_{\mathbb{R}^n}.$$

Nel caso in cui si voglia soltanto sapere se la matrice $A \in \mathcal{M}_n(\mathbb{R})$ è diagonalizzabile o meno, basterà verificare se per ogni autovalore α_i di molteplicità algebrica n_i si ha $\dim(V_{\alpha_i}) = n - \mathrm{rg}(A - \alpha_i I_n) = n_i$. (Si noti che per gli autovalori con molteplicità algebrica 1 tale uguaglianza è sempre verificata). In caso affermativo la matrice è diagonalizzabile, altrimenti non lo è.

Per trovare poi una base di autovettori di A si dovrà scegliere una base di ogni autospazio e prendere l'unione delle basi trovate.

Esempio 9.2.6 Si consideri la matrice

$$T = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Dire se T è diagonalizzabile e, in caso affermativo, trovare una matrice che la diagonalizzi, i.e. una matrice $H \in GL_3(\mathbb{R})$ per cui $H^{-1}TH$ sia una matrice diagonale.

Svolgimento. Seguiamo esattamente le linee guida del procedimento che abbiamo illustrato. Calcoliamo dunque il polinomio caratteristico di T nella incognita λ:

$$\det(\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} - \lambda I_3) = \det \begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{pmatrix} = P_T(\lambda).$$

Sviluppando il determinante di $T - \lambda I_3$ rispetto alla prima riga e ricordando che il nostro obbiettivo è calcolare le radici del polinomio caratteristico di T, abbiamo: $P_T(\lambda) = (3-\lambda)(-\lambda(3-\lambda) - 4) - 2(-2\lambda - 2) + 4(4 + 4\lambda) = (\lambda+1)((3-\lambda)(\lambda-4)+20) = -(\lambda+1)^2(\lambda-8)$. Il polinomio caratteristico di T è dunque fattorizzabile in $\mathbb{R}[\lambda]$ in fattori di primo grado. Le sue radici sono: 8, di molteplicità algebrica 1, e -1 di molteplicità algebrica 2. Per verificare che la matrice T sia diagonalizzabile occorre allora controllare che l'autospazio relativo all'autovalore -1 abbia dimensione due, cioè che la dimensione di V_{-1} coincida con la molteplicità algebrica dell'autovalore -1. Per l'autovalore 8 il risultato è vero automaticamente dal momento che esso ha molteplicità algebrica uguale ad 1.

Determiniamo l'autospazio relativo all' autovalore -1. Tale autospazio è dato dalle soluzioni del sistema:

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (-1) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

cioè dall'insieme delle terne (x_1, x_2, x_3) che soddisfano il sistema

$$(T + I_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3+1 & 2 & 4 \\ 2 & 0+1 & 2 \\ 4 & 2 & 3+1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

A priori sappiamo che la dimensione di V_{-1}, cioè dello spazio delle soluzioni che stiamo cercando, è uguale ad 1 o a 2. Per calcolare tale dimensione osserviamo che le colonne della matrice $T + I_3$ sono una multipla dell'altra, pertanto il rango della matrice $T + I_3$ è uguale a 1. Di conseguenza $\dim(V_{-1}) = 3 - 1 = 2$. Quindi la matrice T è pertanto diagonalizzabile

Determiniamo ora una base di autovettori di T. Cominciamo con l'autospazio relativo all'autovalore -1: dal momento che la matrice $T + I_3$ ha rango 1, V_{-1} è dato dalle soluzioni di una sola equazione: $2x_1 + x_2 + 2x_3 = 0$. Due soluzioni linearmente indipendenti di questa equazione sono, ad esempio, $(1, -2, 0)$ e $(0, -2, 1)$. Dunque $V_{-1} = \langle (1, -2, 0), (0, -2, 1) \rangle$.

Calcoliamo ora l'autospazio relativo all'autovalore 8. Il sistema che caratterizza tale autospazio è allora:

$$(T - 8I_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3-8 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & 3-8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Sappiamo già che l'insieme delle soluzioni di questo sistema è uno spazio di dimensione 1 (la molteplicità dell'autovalore 8 come radice del polinomio caratteristico è infatti uguale ad 1 e la dimensione di V_8, che è certo non banale, non può essere più grande di 1). Quindi V_8 è l'insieme delle soluzioni di due equazioni linearmente indipendenti, ad esempio quelle individuate dalla prima e dalla seconda riga della matrice $T - 8I_3$. Per sostituzione otteniamo: $x_1 = 4x_2 - x_3$ e $-20x_2 + 5x_3 + 2x_2 + 4x_3 = 0$, cioè $x_3 = 2x_2$ e $x_1 = 2x_2$. Si ha pertanto $V_8 = \langle (2, 1, 2) \rangle$. Una base di autovettori di T è quindi data da $v_1 = (1, -2, 0), v_2 = (0, -2, 1), v_3 = (2, 1, 2)$ e la matrice di passaggio da questa base a quella di partenza è la matrice

$$H = \begin{pmatrix} 1 & 0 & 2 \\ -2 & -2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

In conclusione la forma diagonale della matrice T è

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix} = H^{-1}TH.$$

Osserviamo che la matrice D è la matrice dell'applicazione lineare di partenza (associata alla matrice T rispetto alla base canonica di $\mathbb{R}^3$) rispetto alla base v_1, v_2, v_3. Le sue colonne sono dunque date dalle coordinate dei vettori $T(v_1)$, $T(v_2)$, $T(v_3)$ rispetto alla base v_1, v_2, v_3. Poiché questi vettori sono autovettori si ha, effettivamente, $T(v_1) = -v_1 = (-1)v_1 + 0v_2 + 0v_3$, $T(v_2) = -v_2 = 0v_1 + (-1)v_2 + 0v_3$, $T(v_3) = 8v_3 = 0v_1 + 0v_2 + 8v_3$.

Esempio 9.2.7 Stabilire, al variare di $t \in \mathbb{R}$, se la matrice

$$B_t = \begin{pmatrix} t-1 & t-3 & t \\ 0 & \frac{5}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2} \end{pmatrix}$$

è diagonalizzabile o meno.

Svolgimento. Calcoliamo innanzitutto il polinomio caratteristico della matrice B_t:

$$\det(\begin{pmatrix} t-1-\lambda & t-3 & t \\ 0 & \frac{5}{2}-\lambda & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-\lambda \end{pmatrix}) = (t-1-\lambda)(\lambda^2 - 5\lambda + 6).$$

Le sue radici sono : $t-1, 2, 3$. Ora se $t \neq 3, 4$ le tre radici sono diverse e quindi la matrice è diagonalizzabile.

Si noti che per ogni $t \neq 3, 4$ l'autospazio relativo all'autovalore $t-1$ è in ogni caso $\langle(1, 0, 0)\rangle$.

Consideriamo ora i casi $t = 3, 4$: in questi casi dobbiamo studiare direttamente le matrici B_3 e B_4. La matrice B_3 ha autovalore 2 di molteplicità algebrica due e quindi, affinché essa sia diagonalizzabile, occorre che l'autospazio

V_2 relativo all'autovalore 2 abbia dimensione 2. Si ha:

$$\dim(V_2) = 3-\mathrm{rg}(B_3-2I_3) = 3-\mathrm{rg}(\begin{pmatrix} 3-1-2 & 3-3 & 3 \\ 0 & \frac{5}{2}-2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-2 \end{pmatrix}) = 3-2 = 1.$$

Quindi la dimensione di V_2 non coincide con la molteplicità algebrica dell'autovalore 2: la matrice B_3 non è diagonalizzabile.

Consideriamo la matrice B_4: essa ha autovalore 3 di molteplicità algebrica 2. Calcoliamo la dimensione di V_3:

$$\dim(V_3) = 3-\mathrm{rg}(B_4-3I_3) = 3-\mathrm{rg}(\begin{pmatrix} 4-1-3 & 4-3 & 4 \\ 0 & \frac{5}{2}-3 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{2}-3 \end{pmatrix}) = 3-2 = 1.$$

Anche in questo caso la dimensione dell'autospazio V_3 non coincide con la molteplicità algebrica dell'autovalore 3. La matrice B_4 non è diagonalizzabile.

Esempio 9.2.8 Si determinino gli autovalori e gli autospazi della matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -2 & 0 & 1 \\ -2 & 0 & -1 & 2 \end{pmatrix}.$$

Svolgimento. Come al solito occorre calcolare il polinomio caratteristico della matrice A:

$$\det(\begin{pmatrix} 1-\lambda & 0 & 0 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ -2 & -2 & -\lambda & 1 \\ -2 & 0 & -1 & 2-\lambda \end{pmatrix}) = (1-\lambda)^2(\lambda^2 - 2\lambda + 1) = (1-\lambda)^4.$$

Dunque la matrice A ha un unico autovalore $\lambda = 1$ di molteplicità algebrica 4. Studiamo il relativo autospazio. Dobbiamo cercare le soluzioni del sistema

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -2 & 0 & 1 \\ -2 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

cioè le soluzioni del sistema

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2 & -2 & -1 & 1 \\ -2 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

La matrice del sistema omogeneo trovato ha rango due e quindi l'autospazio relativo all'autovalore 1 ha dimensione 2, in particolare la matrice A non è diagonalizzabile. L'autospazio V_1 è l'insieme delle soluzioni del sistema

$$\begin{cases} -2x_1 - 2x_2 - x_3 + x_4 = 0 \\ -2x_1 - x_3 + x_4 = 0 \end{cases}$$

da cui otteniamo: $x_3 = -2x_1 + x_4$ e $x_2 = 0$. Una base di V_1 è allora data dai vettori $(1, 0, -2, 0)$ e $(0, 0, 1, 1)$: $V_1 = \langle (1, 0, -2, 0), (0, 0, 1, 1) \rangle$.

Si noti che se la matrice A fosse stata diagonalizzabile la sua forma diagonale sarebbe stata la matrice identica I_4. In questo caso avremmo avuto $H^{-1}AH = I_4$ per qualche matrice invertibile H di ordine 4. Allora, moltiplicando entrambi i membri dell'uguaglianza a destra per H^{-1} e a sinistra per H, avremmo avuto: $H(H^{-1}AH)H^{-1} = HI_4H^{-1}$ cioè $A = I_4$. Ma questo è ovviamente falso! Si noti ancora che se una matrice $B \in \mathcal{M}_n(\mathbb{R})$ è diagonalizzabile e il suo polinomio caratteristico è $(-1)^n(\lambda - \alpha)^n$, con $\alpha \in \mathbb{R}$, allora $B = \alpha I_n$, dove I_n è la matrice identica in $\mathcal{M}_n(\mathbb{R})$.

Risoluzione problema B descritto nell'introduzione. Descriviamo in termini matriciali il problema B proposto nella introduzione. La situazione del nostro sistema di volpi e galline dopo un anno di osservazione può essere descritto come segue:

$$\begin{pmatrix} G_1 \\ V_1 \end{pmatrix} = \begin{pmatrix} 1,6 & -1 \\ -0,5 & 1,1 \end{pmatrix} \begin{pmatrix} G_0 \\ V_0 \end{pmatrix} = A \begin{pmatrix} G_0 \\ V_0 \end{pmatrix},$$

dove, lo ricordiamo, G_0 e V_0 sono, rispettivamente, il numero di galline ed il numero di volpi al momento iniziale di osservazione del sistema. Dopo n anni la situazione del sistema sarà descritta dalla seguente relazione matriciale:

$$\begin{pmatrix} G_n \\ V_n \end{pmatrix} = \begin{pmatrix} 1,6 & -1 \\ -0,5 & 1,1 \end{pmatrix}^n \begin{pmatrix} G_0 \\ V_0 \end{pmatrix}.$$

D'altro canto se studiamo il polinomio caratteristico della matrice A, troviamo che essa ammette due autovalori: $0,6$ e $2,1$ e che la matrice che diagonalizza A (A è certamente diagonalizzabile poiché ha due autovalori diversi) è

$$H = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}.$$

Si noti che

$$H^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

ed $A = H^{-1} \begin{pmatrix} 0,6 & 0 \\ 0 & 2,1 \end{pmatrix} H$. Osserviamo che

$$A^2 = (H^{-1} \begin{pmatrix} 0,6 & 0 \\ 0 & 2,1 \end{pmatrix} H)(H^{-1} \begin{pmatrix} 0,6 & 0 \\ 0 & 2,1 \end{pmatrix} H) =$$

$$= H^{-1} \begin{pmatrix} 0,6 & 0 \\ 0 & 2,1 \end{pmatrix}^2 H = H^{-1} \begin{pmatrix} (0,6)^2 & 0 \\ 0 & (2,1)^2 \end{pmatrix} H.$$

Ne segue che, scelte le quantità iniziali G_0 e V_0, si ha

$$\begin{pmatrix} G_n \\ V_n \end{pmatrix} = H^{-1} \begin{pmatrix} (0,6)^n & 0 \\ 0 & (2,1)^n \end{pmatrix} H \begin{pmatrix} G_0 \\ V_0 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} (0,6)^n & 0 \\ 0 & (2,1)^n \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} G_0 \\ V_0 \end{pmatrix}$$

i.e.

$$\begin{pmatrix} G_n \\ V_n \end{pmatrix} = (0,6)^n(\frac{1}{3}G_0 + \frac{2}{3}V_0) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (2,1)^n(G_0 - V_0) \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}.$$

Si vede allora che, per qualsiasi scelta di G_0 e V_0, quando n cresce il primo addendo della somma trovata diventa piccolo ($0,6 < 1$!). Dunque se siamo interessati a capire cosa succede al nostro sistema col passare del tempo possiamo trascurare il primo addendo. Per quanto riguarda il secondo addendo, osserviamo che se scegliamo $G_0 = V_0$ allora esso è nullo quindi, se il numero iniziale di volpi coincide con il numero iniziale di galline, con l'andare del tempo sia le volpi che le galline tenderanno ad estinguersi.

Se scegliessimo un numero iniziale di galline diverso dal numero iniziale di volpi avremmo in ogni caso, col passare del tempo, un andamento diverso perché una fra le quantità G_n e V_n diventerebbe negativa per n molto grande, i.e. dopo un grande numero di anni. Più precisamente, se $G_0 > V_0$ allora resterebbero solo galline che aumenterebbero sempre più e se $G_0 < V_0$ resterebbero solo volpi e l'isola diventerebbe deserta.

9.3 Esercizi svolti

Esercizio 9.3.1 Dimostrare che 0 è autovalore per la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Svolgimento. Osserviamo che 0 è autovalore per la matrice A se esiste un vettore $v \neq 0_{\mathbb{R}^2}$ tale che $Av = 0v = 0_{\mathbb{R}^2}$, cioè se esiste un vettore $v \neq 0_{\mathbb{R}^2}$ nel nucleo della matrice A. In altre parole la matrice A ammette l'autovalore 0 se e solo se essa è non invertibile. In effetti le righe di A sono linearmente dipendenti (anzi, uguali!) perciò $det(A) = 0$. Dunque A non è invertibile.

Esercizio 9.3.2 Si consideri lo spazio vettoriale $\mathbb{R}^{\leq 3}[x]$ dei polinomi di grado minore o uguale a 3 nella variabile x, a coefficienti reali, e l'operatore di derivazione

$$D : \mathbb{R}^{\leq 3}[x] \longrightarrow \mathbb{R}^{\leq 3}[x]$$

che associa ad ogni polinomio la sua derivata prima. Determinare gli autovalori di D e calcolarne i relativi autospazi. Decidere se D è diagonalizzabile e, in caso affermativo, determinarne la forma diagonale.

Svolgimento. Nell'esercizio 5.4.3 abbiamo determinato la matrice associata all'applicazione lineare D rispetto alla base $\{1, x, x^2, x^3\}$ di $\mathbb{R}^{\leq 3}[x]$ (detta base canonica):

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Il polinomio caratteristico della matrice D è dunque:

$$\det(D - \lambda I_4) = \det \begin{pmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 2 & 0 \\ 0 & 0 & -\lambda & 3 \\ 0 & 0 & 0 & -\lambda \end{pmatrix} = \lambda^4.$$

La matrice D ha pertanto un solo autovalore $\lambda = 0$ di molteplicità algebrica 4. Allora D non è certamente diagonalizzabile: se lo fosse la sua forma diagonale sarebbe la matrice identicamente nulla, ma l'unica matrice simile alla matrice nulla è la matrice nulla, e D non è certamente la matrice nulla!

Determiniamo l'autospazio V_0 relativo all'autovalore $\lambda = 0$: $V_0 = KerD = \{p(x) \in \mathbb{R}^{\leq 3}[x] \mid D(p(x)) = 0\} = \mathbb{R}$. In particolare $\dim(V_0) = 1$.

Esercizio 9.3.3 Calcolare gli autovalori della matrice $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix}$. Stabilire se essa è diagonalizzabile su $\mathbb{R}$ ed in caso affermativo scrivere la forma diagonale ed una matrice diagonalizzante.

Svolgimento. Calcoliamo il polinomio caratteristico della matrice A:

$$\det(A - tI_3) = \det \begin{pmatrix} 2-t & 1 & 0 \\ 1 & 1-t & -1 \\ 0 & -1 & 2-t \end{pmatrix} = (2-t)((1-t)(2-t)-1)-$$

$(2-t) = t(2-t)(t-3)$.

La matrice A ha 3 autovalori distinti: $t_1 = 0$, $t_2 = 2$, $t_3 = 3$ ed è pertanto diagonalizzabile. Esiste cioè una matrice invertibile H tale che $H^{-1}AH = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Determiniamo la matrice H: le sue colonne costituiscono una base di $\mathbb{R}^3$ di autovettori di A. Si tratta dunque di determinare gli autospazi della matrice A. Cominciamo con $V_0 = \mathrm{Ker}A = \{(x,y,z) \mid A(x,y,z)^t = (0,0,0)^t\}$. Facendo i calcoli si ottiene $V_0 = \langle(1,-2,-1)\rangle$.

Analogamente $V_2 = \{(x,y,z) \mid (A-2I_3)(x,y,z)^t = (0,0,0)^t\} = \langle(1,0,1)\rangle$.

Infine, $V_3 = \{(x,y,z) \mid (A-3I_3)(x,y,z)^t = (0,0,0)^t\} = \langle(1,1,-1)\rangle$.

Siamo ora in grado di costruire la matrice H:

$$H = \begin{pmatrix} 1 & 1 & 1 \\ -2 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix}.$$

Osservazione Si noti l'ordine in cui abbiamo scritto le componenti degli autovettori nelle colonne della matrice H: nella prima colonna appaiono le componenti di un elemento v_0 di V_0, nella seconda le componenti di un elemento v_2 di V_2, infine, nella terza colonna, le componenti di un elemento v_3 di V_3. Che cosa sarebbe successo se avessimo disposto i vettori sulle colonne della matrice in un ordine diverso? Naturalmente avremmo trovato un'altra matrice diagonalizzante la matrice di partenza A. Ad esempio, posto $K = (v_2, v_0, v_3)$ (dove con v_i intendiamo i vettori colonna), si ha:

$$K^{-1}AK = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Esercizio 9.3.4 Sia $V = \mathcal{M}_2(\mathbb{R})$ e sia f l'applicazione lineare di V in sé definita da:

$$f(X) = XA \qquad\qquad (X \in V)$$

dove $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$. Determinare gli autovalori di f.

Svolgimento. Fissiamo innanzitutto la base canonica $\mathcal{C}$ di $\mathcal{M}_2(\mathbb{R})$: $\mathcal{C} = \{e_{11}, e_{12}, e_{21}, e_{22}\}$ dove $e_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $e_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Determiniamo ora la matrice F associata all'applicazione lineare f rispetto alla base $\mathcal{C}$ sia nel dominio che nel codominio. Dal momento che lo spazio vettoriale $V = \mathcal{M}_2(\mathbb{R})$ ha dimensione 4, la matrice F sarà una matrice quadrata di ordine 4. Abbiamo:

$f(e_{11}) = e_{11}A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = e_{12}$;

$f(e_{12}) = e_{12}A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = e_{11} + 2e_{12}$;

$f(e_{21}) = e_{21}A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = e_{22}$;

$f(e_{22}) = e_{22}A = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix} = e_{21} + 2e_{22}$.

Otteniamo così la matrice

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

il cui polinomio caratteristico è:

$$\det(F - tI_4) = \det \begin{pmatrix} -t & 1 & 0 & 0 \\ 1 & 2-t & 0 & 0 \\ 0 & 0 & -t & 1 \\ 0 & 0 & 1 & 2-t \end{pmatrix} = (t^2 - 2t - 1)^2.$$

Gli autovalori di F sono le radici del polinomio $t^2 - 2t - 1$ cioè $t_1 = 1 + \sqrt{2}$, $t_2 = 1 - \sqrt{2}$.

Esercizio 9.3.5

Studiare, al variare del parametro reale k, la diagonalizzabilità della matrice $A_k = \begin{pmatrix} 1 & k & 0 \\ 1 & 0 & 0 \\ 1-k & 0 & 1 \end{pmatrix}$.

Svolgimento. Calcoliamo per prima cosa il polinomio caratteristico della matrice A_k:

$$p_k(t) = \det(A_k - tI_3) = \det \begin{pmatrix} 1-t & k & 0 \\ 1 & -t & 0 \\ 1-k & 0 & 1-t \end{pmatrix} = (1-t)(t^2 - t - k).$$

Il polinomio $p_k(t)$ ha radici $t_1 = 1$, $t_2 = \frac{1+\sqrt{1+4k}}{2}$, $t_3 = \frac{1-\sqrt{1+4k}}{2}$. Così se $k \neq 0, -\frac{1}{4}$ le radici di $p_k(t)$ sono distinte e la matrice A_k è pertanto diagonalizzabile.

Sia ora $k = 0$. Allora l'autovalore $t = 1$ ha molteplicità algebrica 2. Calcoliamo la sua molteplicità geometrica (la molteplicità algebrica di un autovalore è la sua molteplicità come radice del polinomio caratteristico; la molteplicità geometrica di un autovalore è la dimensione del corrispondente autospazio):

$$\dim(V_1) = 3 - \mathrm{rg}(A_0 - I_3) = 3 - \mathrm{rg} \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 1.$$

Dal momento che la molteplicità geometrica di $t = 1$ non coincide con la sua molteplicità algebrica, per $k = 0$ la matrice A_k non è diagonalizzabile.

Sia, infine, $k = -\frac{1}{4}$. In questo caso l'autovalore $t = \frac{1}{2}$ ha molteplicità algebrica 2. Si ha:

$$\dim(V_{\frac{1}{2}}) = 3 - \mathrm{rg}(A_{-\frac{1}{4}} - \frac{1}{2}I_3) = 3 - \mathrm{rg} \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 1 & -\frac{1}{2} & 0 \\ \frac{5}{4} & 0 & \frac{1}{2} \end{pmatrix} = 3 - 2 = 1.$$

Anche in questo caso, dunque, la matrice A_k non è diagonalizzabile.

In conclusione A_k è diagonalizzabile se e solo se $k \neq 0, -\frac{1}{4}$.

9.4 Esercizi proposti

Esercizio 9.4.1 Stabilire per quali valori del parametro reale k la matrice $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & k \\ 0 & -k & 0 \end{pmatrix}$ è diagonalizzabile.

Esercizio 9.4.2 Data la matrice $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$:

1. determinare gli autovalori di A e gli autospazi ad essi relativi;
2. stabilire se la matrice A è diagonalizzabile e, in caso affermativo, determinare una forma diagonale della matrice e la relativa matrice diagonalizzante;
3. calcolare A^{40}.

Esercizio 9.4.3 Stabilire per quali valori del parametro reale h la matrice $A = \begin{pmatrix} 1 & 0 & h \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ è diagonalizzabile. Per ognuno dei valori trovati determinare una base di $\mathbb{R}^3$ costituita da autovettori di A.

Esercizio 9.4.4 Stabilire se esistono valori del parametro reale k tali che la matrice

$$K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ k & 0 & 0 & 0 \end{pmatrix}$$

sia diagonalizzabile su $\mathbb{R}$. Esistono valori di k tali che K sia diagonalizzabile su $\mathbb{C}$?

Lezione 10

Esercizi di ricapitolazione

In questa sezione vogliamo svolgere una serie di esercizi di ricapitolazione.

Esercizio 10.1 Si considerino i due sottoinsiemi di $\mathbb{R}^4$:

$$V = \{(x, y - z, z, y + z) \mid x, y, z \in \mathbb{R}\}$$

$$W = \{(x, 0, y, 2y) \mid x, y \in \mathbb{R}\}.$$

i) Mostrare che V e W sono sottospazi vettoriali di $\mathbb{R}^4$.

ii) Determinare le dimensioni di V e W, esplicitarne delle basi e mostrare che $W \leq V$.

Svolgimento. I vettori di V sono, al variare di x, y, z nell'insieme dei numeri reali, della forma $(x, y-z, z, y+z)$. Dobbiamo mostrare che se consideriamo due vettori v_1 e v_2 di V allora la loro somma in $\mathbb{R}^4$ è ancora un vettore di V, e che se $\alpha \in \mathbb{R}$ allora per ogni $v \in V$, $\alpha v \in V$. Descriviamo allora i vettori v_1 e v_2: poiché stanno in V, esisteranno $x_1, y_1, z_1 \in \mathbb{R}$ tali che $v_1 = (x_1, y_1-z_1, z_1, y_1+z_1)$ e $x_2, y_2, z_2 \in \mathbb{R}$ tali che $v_2 = (x_2, y_2-z_2, z_2, y_2+z_2)$, quello che si vuole verificare è che $v_1 + v_2 \in V$ cioè che esistono $x_3, y_3, z_3 \in \mathbb{R}$ tali che $v_1 + v_2 = (x_3, y_3 - z_3, z_3, y_3 + z_3)$. Per definizione di somma in $\mathbb{R}^3$ si ha: $v_1 + v_2 = (x_1, y_1 - z_1, z_1, y_1 + z_1) + (x_2, y_2 - z_2, z_2, y_2 + z_2) = (x_1 + x_2, y_1 - z_1 + y_2 - z_2, z_1 + z_2, y_1 + z_1 + y_2 + z_2) = ((x_1 + x_2), (y_1 + y_2) - (z_1+z_2), z_1+z_2, (y_1+y_2)+(z_1+z_2))$, quindi avremo la scrittura voluta ponendo $x_3 = x_1+x_2, y_3 = y_1+y_2, z_3 = z_1+z_2$, ottenendo $v_1+v_2 \in V$. Ora prendiamo $v = (x, y-z, z, y+z) \in V$ e $\alpha \in \mathbb{R}$. Allora $\alpha v = (\alpha x, \alpha(y-z), \alpha z, \alpha(y+z))$. Posti $x_1 = \alpha x, y_1 = \alpha y, z_1 = \alpha z \in \mathbb{R}$, si ha $\alpha v = (x_1, y_1 - z_1, z_1, y_1 + z_1)$

quindi $\alpha v \in V$. Si dimostra analogamente che W è un sottospazio vettoriale di $\mathbb{R}^4$.

Passiamo alla seconda parte dell'esercizio. Si vede subito che ogni vettore di V si scrive come $x(1,0,0,0) + y(0,1,0,1) + z(0,-1,1,1)$ cioè come combinazione lineare dei 3 vettori $(1,0,0,0), (0,1,0,1), (0,-1,1,1)$. Dunque i vettori $(1,0,0,0)$, $(0,1,0,1)$, $(0,-1,1,1)$ generano V. Essi sono inoltre linearmente indipendenti infatti, scrivendo una loro combinazione lineare per cui si abbia $x(1,0,0,0) + y(0,1,0,1) + z(0,-1,1,1) = (0,0,0,0)$ si ha necessariamente $x = 0, y - z = 0, z = 0, y + z = 0$ cioè $x = y = z = 0$. Quindi i vettori $(1,0,0,0)$, $(0,1,0,1)$, $(0,-1,1,1)$ individuano una base di V e V ha dimensione 3. Nello stesso modo osserviamo che i vettori di W sono combinazioni lineari di $(1,0,0,0)$ e $(0,0,1,2)$ e i due vettori $(1,0,0,0), (0,0,1,2)$ (che sono quindi dei generatori di W) sono linearmente indipendenti. Infatti da $x(1,0,0,0) + y(0,0,1,2) = (0,0,0,0)$ si ottiene $x = y = 0$. Dunque abbiamo trovato una base di W e possiamo affermare che W ha dimensione 2. Per mostare adesso che $W \leq V$ basterà osservare che ogni elemento della base di W appena trovata è un vettore di V, perché allora ogni combinazione lineare dei vettori della base (cioè ogni vettore di W) starà in V. Si vede che $(1,0,0,0) = 1(1,0,0,0) + 0(0,1,0,1) + 0(0,-1,1,1)$ appartiene a V e anche $(0,0,1,2) = 0(1,0,0,0) + 1(0,1,0,1) + 1(0,-1,1,1)$ appartiene a V. Quindi W è un sottospazio di V.

Esercizio 10.2 Sia V uno spazio vettoriale reale di dimensione 3 e sia v_1, v_2, v_3 una sua base.

i) Esiste un endomorfismo ϕ di V tale che $\phi(v_1) = 4v_1 - 8v_2$, $\phi(v_2) = v_1 + v_2$ e $\phi(v_1 - 3v_2) = v_1 - 10v_2$? In caso affermativo descriverli tutti.

ii) Determinare un endomorfismo φ (i.e. dare una sua matrice rispetto ad una appropriata base) di V tale che $\varphi(v_1) = v_1 - v_2$, $\varphi(v_1 + 2v_2) = v_1 + v_2$ e tale che il suo nucleo contenga $v_3 - v_1$. Un siffatto endomorfismo è unico?

iii) Esiste un endomorfismo f di V tale che $f(v_1) = v_1 - v_2$, $f(v_2) = 3v_1 + v_2$ e $f(v_1 - v_2) = -2v_1 - 2v_2$? In caso affermativo descriverli tutti e in ogni caso giustificare la risposta.

Svolgimento. *i*) Nel primo quesito ci vengono fornite le immagini dei vettori v_1 e v_2: $\phi(v_1)$ e $\phi(v_2)$ e l'immagine del vettore $v_1 - 3v_2$. Ma se esistesse una applicazione lineare come richiesta si avrebbe $\phi(v_1 - 3v_2) = \phi(v_1) - 3\phi(v_2)$, quindi avremmo:

$$v_1 - 10v_2 = \phi(v_1 - 3v_2) = \phi(v_1) - 3\phi(v_2) = (4v_1 - 8v_2) - 3(v_1 + v_2) = v_1 - 11v_2$$

ma questa non è una uguaglianza! Dunque è impossibile che esista una siffatta applicazione lineare.

ii) In questo secondo quesito i dati coinvolgono 3 vettori: vengono fornite le immagini di v_1, v_1+2v_2, v_3-v_1 mediante φ: $\varphi(v_1) = v_1 - v_2$, $\varphi(v_1+2v_2) = v_1 + v_2$, $\varphi(v_3 - v_1) = \mathbf{0}_V$. Se i tre vettori di cui abbiamo le immagini sono linearmente indipendenti si ha che l'applicazione lineare è univocamente determinata: ogni vettore del dominio si scriverà in modo unico come combinazione lineare dei vettori della base e, sfruttando la linearità di φ, si conoscerà l'immagine di ogni vettore di V. Osserviamo che i vettori v_1, v_1+2v_2, v_3-v_1 sono linearmente indipendenti: infatti consideriamo una loro combinazione lineare che dia il vettore nullo: $\alpha_1 v_1+\alpha_2(v_1+2v_2)+\alpha_3(v_3-v_1) = \mathbf{0}_V$, allora: $(\alpha_1+\alpha_2-\alpha_3)v_1+2\alpha_2 v_2+\alpha_3 v_3 = \mathbf{0}_V$, ed essendo i vettori v_1, v_2, v_3 linearmente indipendenti ogni coefficiente è uguale a zero: $\alpha_3 = \alpha_2 = \alpha_1 = 0$. I tre vettori sono linearmente indipendenti in uno spazio di dimensione 3 perciò ne individuano una base. Allora, volendo determinare, ad esempio, l'immagine del vettore $v_1 + 2v_2 - v_3$ basterà scriverlo come combinazione lineare dei vettori v_1, $v_1 + 2v_2$, $v_3 - v_1$ e poi usare la linearità di φ: $v_1 + 2v_2 - v_3 = (\alpha_1 + \alpha_2 - \alpha_3)v_1 + 2\alpha_2 v_2 + \alpha_3 v_3$ con $\alpha_3 = -1$, $\alpha_2 = 1$ e $\alpha_1 = -1$, pertanto $\varphi(v_1 + 2v_2 - v_3) = \varphi(-1(v_1) + 1(v_1 + 2v_2) - 1(v_3 - v_1)) = -1\varphi(v_1)+\varphi(v_1+2v_2)-1\varphi(v_3-v_1) = -1(v_1-v_2)+1(v_1+v_2)-1(\mathbf{0}_V) = 2v_2$. Nello stesso modo si può calcolare l'immagine mediante φ di qualsiasi altro vettore di V.

Dobbiamo ora scrivere la matrice associata alla applicazione lineare rispetto ad una sua base. Conosciamo già le immagini dei vettori della base $w_1 = v_1$, $w_2 = v_1+2v_2$, $w_3 = v_3-v_1$, però i vettori immagine non sono espressi in coordinate rispetto a tale base. Per determinare la matrice rispetto a tale base dobbiamo allora esprimere le immagini dei vettori w_i come combinazioni lineari dei vettori w_1, w_2, w_3. Si ha: $\varphi(w_1) = v_1 - v_2 = \beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3 = \beta_1(v_1) + \beta_2(v_1 + 2v_2) + \beta_3(v_3 - v_1) = (\beta_1 + \beta_2 - \beta_3)v_1 + 2\beta_2 v_2 + \beta_3 v_3$ cioè $\beta_3 = 0$, $\beta_2 = -\frac{1}{2}$ e $\beta_1 = \frac{3}{2}$; $\varphi(w_2) = v_1 + v_2 = \beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3 = \beta_1(v_1) + \beta_2(v_1 + 2v_2) + \beta_3(v_3 - v_1) = (\beta_1 + \beta_2 - \beta_3)v_1 + 2\beta_2 v_2 + \beta_3 v_3$, cioè $\beta_3 = 0$, $\beta_2 = \frac{1}{2}$ e $\beta_1 = \frac{1}{2}$; infine $\varphi(w_3) = \mathbf{0}_V$. Allora la matrice della applicazione lineare φ rispetto alla base w_1, w_2, w_3 è la seguente:

$$A = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Se adesso volessimo scrivere la matrice della applicazione lineare rispetto alla base v_1, v_2, v_3, basterebbe osservare che la matrice di passaggio dalla base w_1, w_2, w_3 alla base v_1, v_2, v_3 è data da

$$T = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(gli elementi sulle colonne di T sono le coordinate nella base v_1, v_2, v_3 dei vettori w_1, w_2, w_3). La matrice del cambiamento di base dalla base v_1, v_2, v_3 alla base w_1, w_2, w_3 è la matrice T^{-1}:

$$T^{-1} = \begin{pmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Allora la matrice della applicazione lineare nella base v_1, v_2, v_3 è :

$$TAT^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{3}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

iii) Nel terzo quesito abbiamo delle informazioni simili a quelle forniteci nel primo. In questo caso però le condizioni sono compatibili. In effetti se tale f lineare esistesse avremmo che: $-2v_1 - 2v_2 = f(v_1 - v_2) = f(v_1) - f(v_2) = (v_1 - v_2) - (3v_1 + v_2) = -2v_1 - 2v_2$. Ne segue che quello che abbiamo sono solo informazioni sulle immagini di v_1 e v_2. Per determinare completamente l'applicazione f occorre fissare, ad esempio, l'immagine di v_3 (o, comunque, l'immagine di un vettore che assieme a v_1 e v_2 formi una base di V). Sulla immagine di v_3 non abbiamo alcuna condizione prefissata. Quindi la matrice di una applicazione f che soddisfi le richieste dell'esercizio, rispetto alla base $\{v_1, v_2, v_3\}$, è:

$$\begin{pmatrix} 1 & -3 & \alpha \\ -1 & 1 & \beta \\ 0 & 0 & \gamma \end{pmatrix}$$

con $\alpha, \beta, \gamma \in \mathbb{R}$, dove abbiamo posto $f(v_3) = (\alpha, \beta, \gamma)$. Le matrici di questa forma, al variare di α, β, γ in $\mathbb{R}$ sono tutte e sole le matrici che rappresentano le applicazioni lineari richieste dal problema (nella base $\{v_1, v_2, v_3\}$).

Esercizio 10.3 Si consideri lo spazio vettoriale reale $\mathcal{M}_2(\mathbb{R})$ delle matrici quadrate di ordine 2 ad entrate reali.

i) Si mostri che l'insieme delle matrici simmetriche $V \subset \mathcal{M}_2(\mathbb{R})$ è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$ e se ne calcoli la dimensione.

ii) Sia $A \in \mathcal{M}_2(\mathbb{R})$,

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

e si consideri l'applicazione ϕ di V in se stesso data da $\phi(S) = A^t SA$, per ogni $S \in V$. Mostrare che ϕ è una applicazione lineare. Determinarne il nucleo e l'immagine. Trovare i suoi autovalori e autovettori. L'applicazione ϕ è diagonalizzabile?

Svolgimento. i) (Vedi 2.4.4) Si ricordi che lo spazio ambiente ha dimensione 4 e che una sua base è data dalle matrici $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Ora le matrici simmetriche di $\mathcal{M}_2(\mathbb{R})$ sono tutte e sole le matrici quadrate di ordine 2 della forma:

$$\begin{pmatrix} x & y \\ y & z \end{pmatrix}$$

al variare di $x, y, z \in \mathbb{R}$. Siano ora $S_1,\ S_2 \in V$ due matrici simmetriche: $S_1 = \begin{pmatrix} x_1 & y_1 \\ y_1 & z_1 \end{pmatrix}$ e $S_2 = \begin{pmatrix} x_2 & y_2 \\ y_2 & z_2 \end{pmatrix}$, allora

$$S_1 +_{\mathcal{M}_2(\mathbb{R})} S_2 = \begin{pmatrix} x_1 & y_1 \\ y_1 & z_1 \end{pmatrix} +_{\mathcal{M}_2(\mathbb{R})} \begin{pmatrix} x_2 & y_2 \\ y_2 & z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 & y_1 + y_2 \\ y_1 + y_2 & z_1 + z_2 \end{pmatrix}$$

che è ancora una matrice simmetrica perché l'entrata di posto $1, 2$ è uguale alla entrata di posto $2, 1$ (cioè essa coincide con la sua trasposta). Analogamente se $\alpha \in \mathbb{R}$ e S è una matrice simmetrica $\begin{pmatrix} x & y \\ y & z \end{pmatrix}$, si ha

$$\alpha S = \alpha \begin{pmatrix} x & y \\ y & z \end{pmatrix} = \begin{pmatrix} \alpha x & \alpha y \\ \alpha y & \alpha z \end{pmatrix},$$

che è ancora una matrice simmetrica. Quindi l'insieme delle matrici simmetriche di ordine 2 è un sottospazio vettoriale di $\mathcal{M}_2(\mathbb{R})$. Ovviamente esso non può avere dimensione 4, poiché in tal caso coinciderebbe con lo spazio vettoriale ambiente delle matrici quadrate di ordine 2, mentre esistono infinite matrici non simmetriche ad esempio la matrice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Dunque V può avere al massimo dimensione 3 e, in effetti, ha proprio dimensione 3: basta mostrare 3 matrici simmetriche (in V) che siano linearmente indipendenti, ad esempio le matrici

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; M_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \; M_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Le matrici M_1, M_2, M_3 sono simmetriche e linearmente indipendenti come vettori dello spazio vettoriale $\mathcal{M}_2(\mathbb{R})$. Consideriamo infatti una combinazione lineare di M_1, M_2, M_3 uguale al vettore nullo di $\mathcal{M}_2(\mathbb{R})$, cioè alla matrice nulla:

$$\alpha \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

con α, β, γ numeri reali, allora $\alpha = \beta = \gamma = 0$. D'altro canto ogni matrice simmetrica si scrive come combinazione lineare di M_1, M_2, M_3:

$$\begin{pmatrix} x & y \\ y & z \end{pmatrix} = x \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + z \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + y \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

quindi $\dim V = 3$.

ii) Se prendiamo una matrice simmetrica S e una matrice A, ambedue in $\mathcal{M}_2(\mathbb{R})$, allora il prodotto A^tSA è una matrice simmetrica. Si ha infatti $(A^tSA)^t = A^tS^t(A^t)^t = A^tSA$ (perché S è simmetrica e dunque $S^t = S$). Pertanto ϕ è un'applicazione di V in V:

$$\begin{array}{rccl} \phi: & V & \longrightarrow & V \\ & S & \longmapsto & \phi(S) = A^tSA. \end{array}$$

La funzione ϕ è lineare, infatti, prese due matrici simmetriche S_1, S_2, $\phi(S_1 + S_2) = A^t(S_1 + S_2)A$, ma la moltiplicazione righe per colonne è distributiva rispetto alla somma di matrici, quindi $\phi(S_1 + S_2) = A^t(S_1 + S_2)A = A^tS_1A + A^tS_2A = \phi(S_1) + \phi(S_2)$. Se poi prendiamo αS, con S matrice simmetrica e α numero reale, abbiamo $\phi(\alpha S) = A^t(\alpha S)A = \alpha(A^tSA) = \alpha\phi(S)$. Per determinare il nucleo e l'immagine di ϕ è conveniente associare a ϕ una matrice dopo aver scelto una base di V. Consideriamo la base di V individuata nel punto *i*): $\mathcal{B} = \{M_1, M_2, M_3\}$. La matrice associata all'applicazione ϕ rispetto a questa base avrà sulla prima colonna le coordinate del vettore $\phi(M_1)$ nella base $\mathcal{B}$, nella seconda colonna le coordinate del vettore $\phi(M_2)$ nella base $\mathcal{B}$ e nella terza colonna le coordinate di $\phi(M_3)$. Abbiamo $\phi(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = (1,1,1,)_{\mathcal{B}} = 1\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; $\phi(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = (0,1,0)_{\mathcal{B}}$; infine $\phi(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} = (0,2,1)_{\mathcal{B}}$. La matrice associata a ϕ rispetto alla base $\mathcal{B}$ è quindi

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}.$$

Tale matrice ha rango 3 infatti $\det(F) = 1$, cioè la matrice F, e quindi l'applicazione ϕ, sono invertibili. Pertanto $\mathrm{Ker}(\phi)$ è lo spazio banale e $\mathrm{Im}(\phi) = V$. Calcoliamo ora il polinomio caratteristico della matrice F. Si ha:

$$\det(F) = \det(F - tI_3) = \det \begin{pmatrix} 1-t & 0 & 0 \\ 1 & 1-t & 2 \\ 1 & 0 & 1-t \end{pmatrix} = (1-t)^3.$$

L'endomorfismo ϕ ha dunque un solo autovalore $t = 1$ di molteplicità algebrica 3. Determiniamo l'autospazio V_1 relativo all'autovalore $t = 1$. V_1 è l'insieme delle soluzioni del sistema

$$(F - I_3)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè $V_1 = \langle (0,1,0) \rangle$. Dal momento che la dimensione di V_1 non coincide con la molteplicità algebrica dell'autovalore 1, l'endomorfismo ϕ non è diagonalizzabile.

Esercizio 10.4 Al variare del parametro $\alpha \in \mathbb{R}$ si consideri l'applicazione lineare φ_α di uno spazio vettoriale T di dimensione 3 in se stesso data, rispetto ad una base t_1, t_2, t_3, dalla matrice

$$\begin{pmatrix} 2\alpha - 1 & 2 & -2\alpha \\ 2 & 0 & -1 \\ -(2\alpha + 1) & -2\alpha & 2\alpha + 2 \end{pmatrix}.$$

i) Si determinino gli α per cui φ_α abbia autovalore 0.

ii) Per i valori di α di cui al punto *ii*) si dica se le corrispondenti applicazioni φ_α sono diagonalizzabili.

Svolgimento. Dire che φ_α ha autovalore 0 equivale a dire che esiste un vettore non nullo $t \in T$ tale che $\varphi_\alpha(t) = \mathbf{0}_T$ cioè, essendo φ_α un endomorfismo di T, che φ_α non è un isomorfismo o, equivalentemente, che φ_α ha nucleo non banale. Questo equivale a dire che la matrice di ordine 3 associata a φ_α rispetto ad una qualsiasi base di T ha determinante uguale a zero. Sarà proprio questa ultima caratterizzazione quella che useremo. In effetti l'esercizio ci fornisce sia una base di T che la matrice associata a φ_α rispetto a questa base:

$$\begin{pmatrix} 2\alpha-1 & 2 & -2\alpha \\ 2 & 0 & -1 \\ -(2\alpha+1) & -2\alpha & 2\alpha+2 \end{pmatrix}$$

di determinante (sviluppando ad esempio rispetto alla seconda colonna) $4\alpha^2 - 2\alpha - 6$. Quindi tale determinante è uguale a zero se e solo se $\alpha = -1$ o $\alpha = \frac{3}{2}$. Per tutti gli altri valori di α l'applicazione è invertibile e quindi non ammette autovettori di autovalore 0. Rispondiamo ora al secondo quesito. Abbiamo determinato i valori per cui φ_α ha un autovalore uguale a 0: consideriamo innanzitutto il caso $\alpha = -1$. In questo caso la matrice della applicazione lineare φ_{-1} è data da

$$\begin{pmatrix} -3 & 2 & 2 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix}.$$

Sappiamo a priori che la matrice ha un autovalore uguale a 0, vediamo adesso chi sono gli altri. Calcoliamo allora il polinomio caratteristico della matrice:

$$\begin{aligned} \det(\begin{pmatrix} -3 & 2 & 2 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} - \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}) &= \det \begin{pmatrix} -3-x & 2 & 2 \\ 2 & -x & -1 \\ 1 & 2 & -x \end{pmatrix} \\ &= -x^3 - 3x^2 + 4x. \end{aligned}$$

Gli zeri del polinomio $-x^3 - 3x^2 + 4x = -x(x+4)(x-1)$ sono tutti diversi fra loro e sono $0, -4, 1$: ciascuna radice ha molteplicità uno. Dunque la matrice è diagonalizzabile.

Determiniamo una base di $\mathbb{R}^3$ costituita da autovettori di φ_{-1} (anche se questo non è esplicitamente richiesto dall'esercizio). Si tratta di determinare una base per ogni autospazio e di considerare poi l'unione di tale basi. Cominciamo con V_0: V_0 è l'insieme delle soluzioni del sistema

$$\begin{pmatrix} -3 & 2 & 2 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Sappiamo a priori che lo spazio delle soluzioni di questo sistema ha dimensione 1. Basta quindi determinare una soluzione non nulla per avere una base. Procedendo per sostituzione otteniamo $V_0 = \langle (2, -1, 4) \rangle$.

Ora consideriamo V_1: tale autospazio, di dimensione 1, è dato dalle soluzioni del sistema omogeneo

$$\begin{pmatrix} -3 & 2 & 2 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

che equivale al sistema

$$(\begin{pmatrix} -3 & 2 & 2 \\ 2 & 0 & -1 \\ 1 & 2 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{pmatrix} -4 & 2 & 2 \\ 2 & -1 & -1 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Procedendo ancora una volta per sostituzione si trova, ad esempio, la soluzione $(3, 1, 5)$ e dunque $V_1 = \langle (3, 1, 5) \rangle$.

Infine dobbiamo determinare V_{-4}: esso è l'insieme delle soluzioni del sistema

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & -1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

che è risolto, ad esempio, dal vettore $(2, -1, 0)$. Dunque $V_{-4} = \langle (2, -1, 0) \rangle$.

Per $\alpha = \frac{3}{2}$ il discorso è del tutto analogo. La matrice associata a $\varphi_{\frac{3}{2}}$ è :

$$\begin{pmatrix} 2 & 2 & -3 \\ 2 & 0 & -1 \\ -4 & -3 & 5 \end{pmatrix}$$

e il suo polinomio caratteristico è $-x^3 + 7x^2 + 9x$. Tale polinomio ha tre radici distinte e quindi la matrice è diagonalizzabile.

Esercizio 10.5 In $\mathbb{R}^4$ si consideri lo spazio vettoriale

$$U = \langle (1,0,2,3),(0,0,1,1)\rangle$$

e sia Σ_α, al variare di $\alpha \in \mathbb{R}$, lo spazio delle soluzioni del sistema lineare S_α nelle incognite x, y, z, w:

$$S_\alpha = \begin{cases} 2x + 3y - 4z - w = 0 \\ x + 2y - 3z = 0 \\ \alpha x + z - 2w = 0 \end{cases}.$$

i) Determinare le soluzioni Σ_α di S_α.
ii) Dire per quali valori di α Σ_α è in somma diretta con U.

Svolgimento. *i*) Determiniamo le soluzioni Σ_α di ogni sistema S_α. Per ogni α il sistema è un sistema omogeneo e dunque le sue soluzioni formeranno in ogni caso uno spazio vettoriale. La matrice associata a S_α è (in questo caso non vi è differenza tra il rango della matrice incompleta e il rango di quella completa poiché la seconda differisce dalla prima per una colonna di zeri):

$$A_\alpha = \begin{pmatrix} 2 & 3 & -4 & -1 \\ 1 & 2 & -3 & 0 \\ \alpha & 0 & 1 & -2 \end{pmatrix}.$$

Calcoliamo il suo rango. Possiamo prima scambiare le prime due righe e poi sommare alla seconda la prima moltiplicata per -2. Otteniamo la matrice

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ \alpha & 0 & 1 & -2 \end{pmatrix}.$$

Adesso possiamo sommare all'ultima riga la prima moltiplicata per $-\alpha$. Tale operazione non ha nessuna controindicazione: in effetti α può essere qualsiasi numero e quindi il rango della matrice non viene alterato da questa operazione. (Il discorso sarebbe stato diverso nel caso in cui avessimo diviso per α: in tal caso avremmo dovuto supporre α diverso da zero). Otteniamo dunque la matrice

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & -2\alpha & 1+3\alpha & -2 \end{pmatrix}.$$

Sommiamo all'ultima riga la seconda moltiplicata per -2α. Si ottiene:

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 1-\alpha & -2+2\alpha \end{pmatrix}.$$

Ora tale matrice ha sempre rango 3 fuorché nel caso in cui $\alpha = 1$, caso in cui il rango è 2. Sappiamo già che $\dim(\Sigma_\alpha) = 4 - \mathrm{rg}(A_\alpha)$. Pertanto $\dim \Sigma_\alpha = 1$ se $\alpha \neq 1$ e $\dim \Sigma_\alpha = 2$ se $\alpha = 1$. Determiniamo Σ_α. Nel caso in cui $\alpha \neq 1$ abbiamo che $\alpha - 1 \neq 0$ e quindi possiamo dividere l'ultima riga della matrice per $\alpha - 1$. Ne consegue che, lasciando libero w, (che è l'unica variabile che non è coinvolta dai gradini della forma a scalini) si ha, dall'ultima riga, $z = 2w$ (che quindi non dipende da α) e poi, sostituendo successivamente nella seconda e nella prima riga, si ha: $y = 3w$ e $x = 0$. Otteniamo così che per ogni $\alpha \neq 1$ lo spazio vettoriale delle soluzioni di S_α è $\Sigma_\alpha = \langle (0, 3, 2, 1) \rangle$.

Resta il caso Σ_1. In questo caso la matrice associata al sistema, ridotta a scalini per righe, è:

$$\begin{pmatrix} 1 & 2 & -3 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

cioè il sistema di partenza è equivalente al sistema:

$$\begin{cases} x + 2y - 3z = 0 \\ -y + 2z - w = 0 \end{cases}$$

da cui otteniamo $y = 2z - w$ e $x = -2y + 3z = -2(2z - w) + 3z = -z + 2w$. Abbiamo così due soluzioni (x, y, z, w) linearmente indipendenti: $(-1, 2, 1, 0)$ e $(2, -1, 0, 1)$. Pertanto $\Sigma_1 = \langle (-1, 2, 1, 0), (2, -1, 0, 1) \rangle$.

ii) Dobbiamo vedere per quali valori di α la somma di Σ_α e U è diretta: $\Sigma_\alpha + U = \Sigma_\alpha \oplus U$. Ora questo avviene se e solo se, per definizione, $U \cap \Sigma_\alpha = \{(0, 0, 0, 0)\}$. Si vede subito che, poiché, per $\alpha \neq 1$, $\Sigma_\alpha = \langle (0, 3, 2, 1) \rangle$, l'intersezione $\Sigma_\alpha \cap U$ è banale: tutti i vettori di $U = \langle (1, 0, 2, 3), (0, 0, 1, 1) \rangle$ si scrivono infatti come $(\gamma, 0, 2\gamma + \delta, 3\gamma + \delta)$ al variare di $\gamma, \delta \in \mathbb{R}$ e hanno quindi la seconda componente sempre nulla, mentre per $\alpha \neq 1$ la seconda entrata di tutti i vettori non nulli di Σ_α è diversa da zero. Resta il caso di Σ_1. In questo caso lo spazio delle soluzioni di S_α ha dimensione 2: $\Sigma_1 = \langle (-1, 2, 1, 0), (2, -1, 0, 1) \rangle$ dunque la somma di Σ_1 ed U è diretta se $\dim(U \oplus \Sigma_1) = 2 + 2 = 4$ cioè $U \oplus \Sigma_1 = \mathbb{R}^4$. Questo equivale a dire che l'unione di una base di U e di una base di Σ_1 forma una base di $\mathbb{R}^4$ o, equivalentemente, che la matrice

$$D = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ -1 & 2 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{pmatrix}$$

ha determinante non nullo. Non resta che calcolare $\det(D)$: $\det(D) = 2 \neq 0$. Dunque $U + \Sigma_1 = U \oplus \Sigma_1$.

Esercizio 10.6 Discutere, al variare di $\beta \in \mathbb{R}$, la diagonalizzabilità della matrice

$$A_\beta = \begin{pmatrix} 1 & 0 & 1 \\ 1 & \beta - 1 & 3 \\ 2 & \beta - 1 & 4 \end{pmatrix}.$$

Svolgimento. Che cosa vuol dire "discutere la diagonalizzabilità della matrice *al variare di* $\beta \in \mathbb{R}$?" Vuol dire che dobbiamo essere pronti a rispondere ad una domanda del tipo: "Per $\beta = 5$ la matrice A_β è diagonalizzabile?" Per risolvere tale problema si potrebbe pensare di considerare singolarmente ogni numero e verificare ogni volta se la corrispondente matrice sia o meno diagonalizzabile. Questo metodo puntuale non ci darebbe la soluzione per ogni β ma solo per i valori di β su cui eventualmente ci soffermassimo. Come rispondere alla questione in modo generale? Consideriamo β alla stregua di un qualsiasi numero e calcoliamo il polinomio caratteristico di A_β. Naturalmente tale polinomio caratteristico dipenderà da β:

$$\det\left(\begin{pmatrix} 1 & 0 & 1 \\ 1 & \beta-1 & 3 \\ 2 & \beta-1 & 4 \end{pmatrix} - \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x \end{pmatrix}\right) = \det\begin{pmatrix} 1-x & 0 & 1 \\ 1 & \beta-1-x & 3 \\ 2 & \beta-1 & 4-x \end{pmatrix}$$

$$= -x^3 + (4+\beta)x^2 - 2\beta x = x(-x^2 + (4+\beta)x - 2\beta).$$

Tale polinomio caratteristico ha sempre una radice nulla e le sue altre due radici sono le radici del polinomio $x^2 - (4+\beta)x + 2\beta$ cioè $\frac{4+\beta\pm\sqrt{16+\beta^2}}{2}$. Se le radici del polinomio caratteristico fossero una diversa dall'altra potremmo affermare con sicurezza che la matrice A_β è diagonalizzabile. Per quali valori di β si verifica questa circostanza? Osserviamo che le radici del polinomio $x^2 - (4+\beta)x + 2\beta$ sono distinte per ogni valore di β dal momento che il discriminante del polinomio è sempre maggiore di 0. Si tratta di stabilire per quali valori di β una delle due è uguale a 0. Con calcoli algebrici elementari si trova che la radice nulla può avere molteplicità maggiore di 1 (e uguale a 2) solo nel caso in cui β sia uguale a 0. Questo ci consente di affermare che per ogni $\beta \neq 0$ la matrice A_β è diagonalizzabile.

Che cosa succede quando $\beta = 0$? In questo caso il polinomio caratteristico ha una radice $x = 4$ di molteplicità uno e una radice $x = 0$ di molteplicità 2. La matrice A_0 è dunque diagonalizzabile se e solo se l'autospazio V_0 relativo all'autovalore 0 ha dimensione 2. Calcoliamo la dimensione di questo autospazio:

$$\dim(V_0) = 3 - \mathrm{rg}\begin{pmatrix} 1 & 0 & 1 \\ 1 & -1 & 3 \\ 2 & -1 & 4 \end{pmatrix} = 3 - 2 = 1.$$

Dunque per $\beta = 0$ la matrice A_β non è diagonalizzabile.

Esercizio 10.7 Si considerino, al variare di $\lambda \in \mathbb{R}$, le funzioni $L_\lambda : \mathbb{R}^4 \to \mathbb{R}^3$ espresse rispetto ad una base v_1, v_2, v_3 di $\mathbb{R}^3$ e ad una base w_1, w_2, w_3, w_4 di $\mathbb{R}^4$ dalle matrici

$$\Lambda_\lambda = \begin{pmatrix} -\lambda & \lambda & 0 & \lambda \\ 2 & -1 & 1 & 1 \\ 3\lambda & -2\lambda & \lambda & 0 \end{pmatrix}.$$

i) Esistono $\lambda \in \mathbb{R}$ per cui $\dim \mathrm{Ker}(L_\lambda) > 2$? Se sì determinare $\mathrm{Ker}(L_\lambda)$ per siffatti λ.

ii) Per i $\lambda \in \mathbb{R}$ per cui $\dim \mathrm{Ker}(L_\lambda) \leq 2$ determinare $\mathrm{Ker}(L_\lambda)$.

Svolgimento. *i*) Affinché la dimensione del nucleo di L_λ sia più grande di 2, deve succedere che il rango della matrice sia minore o uguale a 1 (anzi uguale a 1, perché per avere rango uguale a zero, la matrice deve avere tutte le entrate uguali a 0, e la nostra matrice ha sicuramente qualche entrata diversa da zero). Per avere rango uguale ad uno la matrice Λ_λ dovrà avere una sola riga linearmente indipendente e l'unica possibilità è che sia $\lambda = 0$ (non appena $\lambda \neq 0$ le prime due righe di Λ_λ sono linearmente indipendenti. Quindi la risposta al primo quesito è $\lambda = 0$. In questo caso la matrice della applicazione lineare è

$$\Lambda_0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

che ha rango 1 e quindi nucleo di dimensione 3: un vettore in coordinate (x_1, x_2, x_3, x_4) appartiene al nucleo di L_0 se

$$\Lambda_0 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

cioè se $2x_1 - x_2 + x_3 + x_4 = 0$. Dal momento che $\mathrm{Ker}(L_0)$ ha dimensione 3 ci occorrono 3 soluzioni linearmente indipendenti dell'equazione $x_2 = 2x_1 + x_3 + x_4$; possiamo scegliere: $x_1 = 1, x_3 = 0, x_4 = 0$, ottenendo la soluzione $(1, 2, 0, 0)$, e $x_1 = 0, x_3 = 1, x_4 = 0$, ottenendo la soluzione $(0, 1, 1, 0)$ e, infine, $x_1 = 0, x_3 = 0, x_4 = 1$, ottenendo la soluzione $(0, 1, 0, 1)$. Questi tre vettori sono linearmente indipendenti ed appartengono a $\mathrm{Ker}(\Lambda_0)$: $\mathrm{Ker}\Lambda_0 = \langle (0, 1, 1, 0), (1, 2, 0, 0), (0, 1, 0, 1) \rangle$.

ii) Sappiamo già che per $\lambda \neq 0$ $\dim(\mathrm{Ker}L_\lambda) \leq 2$. In questo caso la matrice

$$\Lambda_\lambda = \begin{pmatrix} -\lambda & \lambda & 0 & \lambda \\ 2 & -1 & 1 & 1 \\ 3\lambda & -2\lambda & \lambda & 0 \end{pmatrix}.$$

ha di sicuro rango 2: le prime due righe sono linearmente indipendenti (tra parentesi anche le ultime due sono linearmente indipendenti). Cerchiamo di ridurre la matrice Λ_λ in forma a scalini per righe: moltiplichiamo la seconda riga di Λ_λ per $-\lambda$ e sommiamola alla prima:

$$\begin{pmatrix} -3\lambda & 2\lambda & -\lambda & 0 \\ 2 & -1 & 1 & 1 \\ 3\lambda & -2\lambda & \lambda & 0 \end{pmatrix};$$

adesso sostituiamo la terza riga con la somma della prima riga e dell'ultima:

$$\begin{pmatrix} -3\lambda & 2\lambda & -\lambda & 0 \\ 2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

infine sostituiamo alla prima riga la somma della prima con la seconda moltiplicata per $\frac{3\lambda}{2}$ e scambiamo la prima e la seconda riga della matrice ottenuta:

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 0 & \frac{\lambda}{2} & \frac{\lambda}{2} & \frac{3\lambda}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Otteniamo così che la matrice Λ_λ ha rango uguale a 2 per ogni $\lambda \neq 0$. Le trasformazioni per riga non alterano il nucleo (in effetti equivalgono a cambiamenti di base del codominio). Quindi per determinare $\mathrm{Ker}(L_\lambda)$ possiamo utilizzare la forma a scalini per righe che abbiamo trovato. Poiché stiamo assumendo $\lambda \neq 0$ questo equivale a risolvere il sistema:

$$\begin{cases} 2x_1 - x_2 + x_3 + x_4 = 0 \\ x_2 + x_3 + 3x_4 = 0. \end{cases}$$

$\mathrm{Ker}(L_\lambda)$ è dunque l'insieme delle soluzioni di un sistema omogeneo di rango 2 in 4 incognite cioè uno spazio vettoriale di dimensione 2. Lasciamo libere le variabili x_3 e x_4. Allora, ponendo $x_1 = 1$ e $x_2 = 0$ otteniamo per sostituzione $x_3 = -3$ e $x_4 = 1$; dopodiché se poniamo $x_1 = 0$ e $x_2 = 1$ otteniamo $x_3 = 2$ e $x_4 = -1$. Le soluzioni $(1, 0, -3, 1)$ e $(0, 1, 2, -1)$ così individuate sono linearmente indipendenti e $\mathrm{Ker}(L_\lambda) = \langle (1, 0, -3, 1), (0, 1, 2, -1) \rangle$, per ogni $\lambda \neq 0$.

Esercizio 10.8 Al variare di $\alpha \in \mathbb{R}$ si risolva il sistema nelle incognite x, y, z:

$$\begin{cases} 5x + 2z = 4 \\ x + y - \alpha z = 2 \\ 3x - y = 3. \end{cases}$$

Svolgimento. Ricordiamo che rispondere a questa domanda significa essere pronti a dire: chi sono le soluzioni del sistema dato nel caso in cui si sostituisca ad α il numero 57? Ovviamente cerchiamo un metodo del tutto generale che ci permetta di rispondere senza dover cercare la soluzione per ogni singolo α. La matrice completa associata al sistema è:

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 5 & 0 & 2 & 4 \\ 1 & 1 & -\alpha & 2 \\ 3 & -1 & 0 & 3 \end{array} \right).$$

Riducendo tale matrice in forma a scalini per righe otteniamo la matrice

$$(A|\mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 1 & -\alpha & 2 \\ 0 & -5 & 2+5\alpha & -6 \\ 0 & 0 & 8+5\alpha & -9 \end{array} \right).$$

Si ha dunque $rg(A) = rg(A|\mathbf{b})$ se e solo se $\alpha \neq -\frac{8}{5}$. Per il Teorema di Rouché-Capelli, il sistema ammette soluzioni se e solo se $\alpha \neq -\frac{8}{5}$. In tal caso $rg(A) = rg(A|\mathbf{b}) = 3$ quindi il sistema ammette una sola soluzione che si ottiene risolvendo, per sostituzioni successive dal basso, il sistema di equazioni:

$$\begin{cases} x + y - \alpha z = 2 \\ -5y + (2 + 5\alpha)z = -6 \\ (8 + 5\alpha)z = -9 \end{cases}$$

Lasciamo i calcoli finali al lettore.

Esercizio 10.9 Al variare di λ nei numeri reali si consideri il sistema nelle incognite x, y:

$$\begin{cases} 2x + 4y = \lambda \\ \lambda x - 2y = 1 \\ x + 3y = -1. \end{cases}$$

Per ogni valore di λ se ne determino le soluzioni.

Svolgimento. Scriviamo la matrice completa associata al sistema:

$$\left(\begin{array}{cc|c} 2 & 4 & \lambda \\ \lambda & -2 & 1 \\ 1 & 3 & -1 \end{array}\right)$$

La matrice incompleta ha rango 2 poiché ha due righe linearmente indipendenti e non può avere rango maggiore di 2; affinché il sistema abbia soluzioni la matrice completa non può avere rango tre. La matrice completa è la matrice quadrata di ordine 3:

$$\begin{pmatrix} 2 & 4 & \lambda \\ \lambda & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix},$$

quindi se essa ha determinante non nullo il sistema non ha soluzioni. Abbiamo

$$\det \begin{pmatrix} 2 & 4 & \lambda \\ \lambda & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix} = 3\lambda^2 + 6\lambda + 2,$$

quindi se $\lambda \neq -1 + \frac{1}{\sqrt{3}}$ e $\lambda \neq -1 - \frac{1}{\sqrt{3}}$, la matrice completa ha rango 3 e quindi il sistema non ammette soluzioni. Per $\lambda = -1 \pm \frac{1}{\sqrt{3}}$ la matrice completa del sistema ha rango ≤ 2. Poiché la matrice incompleta ha rango 2, anche la matrice completa ha lo stesso rango. Quindi il sistema ammette soluzioni. Sia $\lambda = -1 + \frac{1}{\sqrt{3}}$. Allora la matrice completa del sistema è la matrice

$$C = \begin{pmatrix} 2 & 4 & -1 + \frac{1}{\sqrt{3}} \\ -1 + \frac{1}{\sqrt{3}} & -2 & 1 \\ 1 & 3 & -1 \end{pmatrix}$$

di rango 2. Ne consegue che una delle righe di C deve essere combinazione lineare delle altre due! Dunque possiamo dire che il sistema è equivalente al sistema

$$\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 + \frac{1}{\sqrt{3}} \\ -1 \end{pmatrix}.$$

E determinando l'inversa della matrice $\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$:

$$\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{3}{2} & -2 \\ -\frac{1}{2} & 1 \end{pmatrix},$$

otteniamo la soluzione:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & -2 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} -1 + \frac{1}{\sqrt{3}} \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} + \frac{1}{2} \\ -\frac{1}{2\sqrt{3}} - \frac{1}{2} \end{pmatrix}.$$

Analogamente si procede per $\lambda = -1 - \frac{1}{\sqrt{3}}$

Esercizio 10.10 Nello spazio vettoriale $\mathbb{R}^3$ sono assegnati, al variare di $\lambda \in \mathbb{R}$, il sottospazio

$$W_\lambda = \langle (1, 1 + \lambda, -1), (2, \lambda - 2, 2 + \lambda) \rangle$$

e l'endomorfismo f_λ la cui matrice, rispetto alla base canonica di $\mathbb{R}^3$, è

$$\begin{pmatrix} \lambda - 4 & -8 & -8 \\ 0 & \lambda + 4 & 0 \\ 2\lambda + 4 & 2\lambda + 4 & 3\lambda + 8 \end{pmatrix}.$$

i) Determinare, al variare di λ, la dimensione di $\mathrm{Ker} f_\lambda$.
ii) Determinare, al variare di λ, la dimensione di W_λ.
iii) Determinare per quali valori di λ si abbia

$$W_\lambda \oplus \mathrm{Ker} f_\lambda.$$

Svolgimento. *i*) Per risolvere il primo quesito dobbiamo studiare il rango della matrice che rappresenta f_λ che in ogni caso è la matrice di un endomorfismo di $\mathbb{R}^3$. Tale endomorfismo sarà un isomorfismo, e quindi $\mathrm{Ker} f_\lambda = \{(0,0,0)\}$, se

$$\det \begin{pmatrix} \lambda-4 & -8 & -8 \\ 0 & \lambda+4 & 0 \\ 2\lambda+4 & 2\lambda+4 & 3\lambda+8 \end{pmatrix} \neq 0$$

ma tale determinante è uguale a $(\lambda+4)(3\lambda^2+12\lambda) = 3\lambda(\lambda+4)^2$, ed è quindi diverso da zero se e solo se $\lambda \neq -4$, $\lambda \neq 0$. Quindi per $\lambda \neq 0, -4$ $\dim(\mathrm{Ker} f_\lambda) = 0$.

Nel caso in cui si abbia $\lambda = 0$ la matrice della applicazione lineare è

$$\begin{pmatrix} -4 & -8 & -8 \\ 0 & 4 & 0 \\ 4 & 4 & 8 \end{pmatrix}$$

e ha rango uguale a 2 perché $\det A = 0$ e il minore quadrato $\begin{pmatrix} 0 & 4 \\ 4 & 4 \end{pmatrix}$ è invertibile. Ne segue che $\dim(\mathrm{Ker} f_0) = 3 - 2 = 1$ e, più precisamente, $\mathrm{Ker} f_0 = \langle (2, 0, -1) \rangle$.

Nel caso in cui sia $\lambda = -4$ la matrice associata a f_λ è la matrice

$$\begin{pmatrix} -8 & -8 & -8 \\ 0 & 0 & 0 \\ -4 & -4 & -4 \end{pmatrix}$$

che ha rango uguale a uno, allora $\dim(\mathrm{Ker} f_{-4}) = 2$.

ii) Lo spazio vettoriale W_λ è generato dai due vettori $(1, 1+\lambda, -1), (2, \lambda-2, 2+\lambda)$. Allora W_λ avrà dimensione uguale al rango della matrice

$$\begin{pmatrix} 1 & 1+\lambda & -1 \\ 2 & \lambda-2 & 2+\lambda \end{pmatrix}$$

(infatti il rango di una matrice è il massimo numero di righe linearmente indipendenti della matrice). Per calcolare tale rango possiamo trovare la forma a scalini con una sola operazione sulle righe:

$$\begin{pmatrix} 1 & 1+\lambda & -1 \\ 0 & -\lambda-4 & 4+\lambda \end{pmatrix}.$$

La matrice ottenuta ha rango 2 se $\lambda \neq -4$ e rango 1 se $\lambda = -4$. Quindi $\dim(W_\lambda) = 2$ se $\lambda \neq -4$ e $\dim(W_{-4}) = 1$.

iii) Per stabilire quando i due spazi sono in somma diretta, analizziamo innanzitutto il problema da un punto di vista qualitativo. Siamo in uno spazio ambiente tridimensionale: nel caso in cui $\lambda \neq -4, 0$ $\mathrm{Ker} f_\lambda$ è il sottospazio banale di $\mathbb{R}^3$ e non c'è nulla da dimostrare. Se $\lambda = 0$ allora $\mathrm{Ker} f_0$ ha dimensione uno e W_0 ha dimensione 2. Se $\mathrm{Ker} f_0$ e W_0 sono in somma diretta, per la formula delle dimensioni deve essere $\dim(W_0 + \mathrm{Ker} f_0) = 3$. Ora $\mathrm{Ker} f_0 = \langle(-2, 0, 1)\rangle$ e $W_0 = \langle(1, 1, -1), (2, -2, 2)\rangle$ pertanto essi sono in somma diretta se i loro generatori sono linearmente indipendenti, cioè

$$\det \begin{pmatrix} 1 & 1 & -1 \\ 2 & -2 & 2 \\ -2 & 0 & 1 \end{pmatrix} \neq 0.$$

In effetti tale determinante è pari a -4, quindi la somma di W_0 e $\mathrm{Ker} f_0$ è diretta, anzi, vale di più:

$$W_0 \oplus \mathrm{Ker} f_0 = \mathbb{R}^3.$$

Resta ora il caso $\lambda = -4$. In questo caso, $\dim(\mathrm{Ker} f_{-4}) = 2$ e si ha $\mathrm{Ker} f_{-4} = \langle(1, -1, 0), (1, 0, -1)\rangle$, mentre $W_{-4} = \langle(1, -3, -1)\rangle$. Come nel caso precedente $\mathrm{Ker} f_{-4}$ e W_{-4} sarebbero in somma diretta se

$$\det \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & -3 & -1 \end{pmatrix}$$

fosse diverso da zero, ed è in effetti questo il caso perché tale determinante è uguale a -3.

Esercizio 10.11 In $\mathbb{R}^3$ si considerino i sottospazi

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid x - 2y = 0 = y + z\}$$

e

$$W_a = \{(x, y, z) \in \mathbb{R}^3 \mid ax + y - z = 0\}$$

al variare di $a \in \mathbb{R}$.

Si dica per quali valori di $a \in \mathbb{R}$ si ha $U \oplus W_a = \mathbb{R}^3$. Per i valori di a tali che la somma $U + W_a$ non sia diretta, descrivere tale somma.

Svolgimento. Il sottospazio U è l'insieme delle soluzioni del sistema omogeneo

$$\begin{cases} x - 2y = 0 \\ y + z = 0 \end{cases}$$

nelle incognite x, y, z. La matrice associata a questo sistema è la matrice

$$\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

che ha rango 2, quindi U è un sottospazio vettoriale di $\mathbb{R}^3$ di dimensione 1. È facile calcolare una soluzione di questo sistema e ottenere: $U = \langle (2, 1, -1) \rangle$.

Il sottospazio W_a è, per ogni scelta di $a \in \mathbb{R}$, l'insieme delle soluzioni di un sistema omogeneo di rango 1 in 3 incognite, pertanto W_a ha dimensione 2. Da un punto di vista dimensionale, dunque, i dati non ci consentono di escludere a priori il fatto che U e W_a siano in somma diretta (se entrambi i sottospazi avessero avuto dimensione 2, allora avremmo potuto immediatamente affermare che la loro somma non poteva essere diretta: lo spazio somma avrebbe avuto dimensione 4 contraddicendo il fatto di essere un sottospazio di $\mathbb{R}^3$).

Dobbiamo studiare l'intersezione di U e W_a. I vettori che stanno nella intersezione di U e W_a debbono soddisfare ambedue i sistemi che caratterizzano i due spazi vettoriali, cioè il sistema:

$$\begin{cases} x - 2y = 0 \\ y + z = 0 \\ ax + y - z = 0 \end{cases}$$

o, in forma matriciale, il sistema:

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 1 \\ a & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Osserviamo che il sistema così individuato ha come unica soluzione la soluzione banale se e solo se il rango della matrice A è massimo, cioè per tutti e soli i valori di a per cui $\mathrm{rg}(A) = 3$. In altre parole i sottospazi U e W_a hanno intersezione banale (e quindi sono in somma diretta) per tutti e soli i valori di a per cui $\det(A) \neq 0$. In questo caso si ha $\dim(U \oplus W_a) = \dim U + \dim W_a = 2 + 1 = 3$, cioè $U \oplus W_a = \mathbb{R}^3$. Calcoliamo il determinante della matrice A: $\det(A) = -2a - 2$. Dunque per ogni $a \neq -1$ i due spazi sono in somma diretta. Se $a = -1$ allora $U \cap W_{-1} = T \neq \{(0,0,0)\}$. Del resto $U \cap W_{-1}$ è un sottospazio vettoriale di U, perciò se esso è non banale coincide necessariamente con U perché U ha dimensione 1. Nello stesso tempo $U \cap W_{-1} = U$ è un sottospazio di W_{-1} quindi, se $a = -1$, U è contenuto in W_{-1} e $U + W_{-1} = W_{-1}$.

Lezione 11

Soluzioni degli esercizi proposti

2.5.3 $k = 0$; $\{(-2, 1, 0, 0), (-8, 0, 0, 1), (0, 0, 1, 0)\}$; il vettore $(0, 0, 0, 1)$ non è combinazione lineare dei precedenti.

2.5.4 S non è un sottospazio vettoriale di $\mathbb{R}[x]$; $T = \mathbb{R}^{\leq 3}[x]$.

2.5.5 $S = \{(x, y) \in \mathbb{R}^2 | x - 3y = 0\} \cup \{(x, y) \in \mathbb{R}^2 \mid 2x - y = 0\}$; il più piccolo sottospazio di $\mathbb{R}^2$ contenente S è $\mathbb{R}^2$ stesso.

3.4.2 La base $\mathcal{B}$ non è unica. Una possibilità è: $\mathcal{B} = \{(1, 1, 1), (1, 1, 0), (0, -1, 0)\}$.

3.4.3 I vettori v_1, v_2, v_3 e v_4 non sono linearmente indipendenti e non generano $\mathbb{R}^3$. Una base di $\langle v_1, v_2, v_3, v_4\rangle$ è: $\{(1, 2, 0), (0, 1, -1)\}$ che si può completare nella base $\{(1, 2, 0), (0, 1, -1), (0, 0, 1)\}$ di $\mathbb{R}^3$.

3.4.4 Una base di S è:

$$\{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\};$$

$$T = \langle\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\rangle.$$

3.4.5 $\mathcal{B} = \{1, x, y, x^2, xy, y^2\}$; $\dim(S) = 6$; le coordinate di $x + y - x^2$ nella base $\mathcal{B}$ sono $(0, 1, 1, -1, 0, 0)$; l'insieme $\{x-y, 1+x-y, 1-xy\}$ può essere completato nella seguente base di V: $\{x-y, 1+x-y, 1-xy, x, x^2, y^2\}$.

4.4.2 $\dim(S) = 3$; $T = \langle(0, 0, 0, 1)\rangle$; non esiste un sottospazio W come richiesto; $V = \langle(1, 1, 1, 1), (0, 0, 0, 1)\rangle$.

4.4.3 $U = \langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \rangle$, $V = \langle \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \rangle$; $U \cap V = \{0\}$;

$U + V = \langle \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \rangle$;

$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$;

$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = (0, 1, 1, 0)_{\mathcal{B}}$.

4.4.4 $\dim(S) = 2$, $\dim(T) = 2$; $S \cap T = \langle -u + w \rangle$; $S + T = V$; $\{-u + w, u, v\}$.

4.4.5 $S \cap T = \langle x + x^2 \rangle$; $S + T = \mathbb{R}^{\leq 2}[x]$; la scrittura non è unica.

5.5.1 Le applicazioni f e g non sono lineari. L'applicazione h lo è.

5.5.2 f_k è lineare per ogni valore di k; la matrice associata a f_k rispetto alle basi fissate è: $\begin{pmatrix} 0 & -1 \\ k & 2k \\ k & 2k \end{pmatrix}$; se $k = 0$, $\mathrm{Ker} f_0 = \langle (1, 1) \rangle$ e $\mathrm{Im} f_0 = \langle (1, 0, 0) \rangle$; se $k \neq 0$, $\mathrm{Ker} f_k = \{(0, 0)\}$ e $\mathrm{Im} f_0 = \langle (1, 0, 0), (0, 1, 1) \rangle$; $(2, 3, 3) \in \mathrm{Im} f_k$ per ogni $k \neq 0$.

5.5.4 La matrice associata ad L rispetto alla base $\{1, x, x^2\}$ di $\mathbb{R}^{\leq 2}[x]$ ed alla base canonica di $\mathbb{R}^3$ è: $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$; $\mathrm{Ker} L = \langle -1 + x + x^2 \rangle$, $\mathrm{Im} L = \langle (1, 1, 0), (1, 0, 1) \rangle$; L non è iniettiva e non è suriettiva.

5.5.5 La matrice associata a f rispetto alle basi fissate è: $\begin{pmatrix} 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$; $\mathrm{Ker} f = \langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rangle$, $\mathrm{Im} f = \mathbb{R}^2$; $f^{-1}(S) = \langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rangle + \mathrm{Ker} f$.

6.5.1 Il sistema non ha soluzioni per $h \neq 0, 1$, ha una sola soluzione: $(\frac{5}{9}, \frac{5}{9}, -\frac{7}{9})$ per $h = 1$, e infinite soluzioni della forma $(2t, t, 5t - 3)$, $t \in \mathbb{R}$, per $h = 0$.

6.5.2 Il sistema non ha soluzioni se $a = 0$ e $c \neq 0$ o se $a \neq 0$ e $c \neq (b - 2)a$; ha infinite soluzioni della forma $(t + b, t, s, 0)$, $t, s \in \mathbb{R}$ se $a = 0 = c$ ed ha infinite soluzioni della forma $(\frac{-a+1}{a}t+b-1, \frac{a+1}{a}t-1, s, t)$, $t, s \in \mathbb{R}$, se $a \neq 0$ e $c = (b-2)a$.

6.5.3 $(\frac{1}{11}, \frac{2}{11}, \frac{1}{22})$.

6.5.4 $\begin{cases} x+z=2 \\ -x+y=1. \end{cases}$

6.5.5 $\mathrm{Ker}A = \langle(-2,-1,1)\rangle$, $\mathrm{Im}A = \langle(0,3,1),(1,0,6)\rangle$, $\mathrm{Ker}A^t = \langle(18,1,-3)\rangle$, $\mathrm{Im}A^t = \langle(0,1,1),(1,0,2)\rangle$; $L^{-1}(k,3,7) = \emptyset$ per ogni $k \neq 1$ e $L^{-1}(1,3,7) = (1,1,0) + \langle(-2,-1,1)\rangle$.

6.5.6 $\langle(-1,1,-1,-1),(-1,0,0,1)\rangle$.

7.4.1 $A = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & -1 \\ 1 & -1 & 0 & -1 \\ 0 & 1 & 1 & 3 \end{pmatrix}$.

7.4.2 $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

7.4.4 $(I+A)^{-1} = I - A + A^2 - A^3 + \ldots + (-1)^{k-1}A^{k-1}$.

7.4.5 Esistono infinite matrici B come quella richiesta, un esempio è $B = \begin{pmatrix} 2 & -3 & 0 \\ 0 & -2 & 1 \end{pmatrix}$; non esistono matrici C soddisfacenti le condizioni 2.

8.6.1 $\det(A) = -10$.

8.6.2 $\det(A) = 0$.

8.6.4 $M_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} -1 & -\frac{2}{3} & -\frac{4}{3} \\ -2 & -1 & -2 \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$.

8.6.5 La matrice richiesta è: $D = \begin{pmatrix} 0 & 0 & 0 & -3 \\ 0 & 0 & -2 & 3 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$; $\det(D) = 0$.

9.4.1 $-1 < k < 1$.

9.4.2 Gli autovalori di A sono 3 ed 1; gli autospazi sono $V_3 = \langle(1,1)\rangle$ e $V_1 = \langle(1,-1)\rangle$; una forma diagonale di A è: $\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$, la relativa matrice diagonalizzante è $H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$; $A^{40} = \begin{pmatrix} \frac{3^{40}+1}{2} & \frac{3^{40}-1}{2} \\ \frac{3^{40}-1}{2} & \frac{3^{40}+1}{2} \end{pmatrix}$.

9.4.3 A è diagonalizzabile per ogni $h > -\frac{1}{4}$. Se $h = 0$ una base di $\mathbb{R}^3$ costituita da autovettori di A è: $\{(1,0,1),(0,1,0),(0,0,1)\}$; se $h > -\frac{1}{4}$ e $h \neq 0$, una base di $\mathbb{R}^3$ costituita da autovettori di A è: $\{(0,1,0), (\frac{1+\sqrt{1+4h}}{2},0,1), (\frac{1-\sqrt{1+4h}}{2},0,1)\}$.

9.4.4 La matrice K non è diagonalizzabile su $\mathbb{R}$ per nessun valore di k; essa è diagonalizzabile su $\mathbb{C}$ per ogni valore di $k \neq 0$.